

## REMARKS

The Office is authorized to charge the fee for a three month extension of time to Deposit Account No. 50-0911. Any fees that may be due in connection with the filing of this paper or with this application may be charged to Deposit Account No. 50-0911. If a Petition for Extension of time is needed, this paper is to be considered such Petition.

A supplemental Information Disclosure Statement accompanies this response.

A change in Power of Attorney for the undersigned is has been filed. A Change in Correspondence Address accompanies this response.

Claims 1, 3-7, 9, 11-16, 45, 48, 51-53, 57-59, 61-63, and 65-78 are pending. Claims 1, 5, 59, 63 and 66 are amended and claims 2 and 56 are cancelled without prejudice or disclaimer. The claims are amended for clarity. For example, the claim is reorganized so that it is clear that recited limitations are not part of the preamble. Further basis for the claim amendments can be found in the specification as originally filed, for example, at page 20, lines 3-7; at page 25, lines 3-16, and at page 28, lines 1-3; and Table 3, which recite particular specificity determinants S1-S4 and methods of mutating the specificity determinants to generate protease muteins.

The arguments and response filed August 7, 2008, June 11, 2009 and July 20, 2010 are each incorporated by reference herein.

### **I. CLAIM OBJECTIONS**

Claim 66 is objected to because it depends upon cancelled claim 64. This objection is addressed by amendment of the claims herein.

### **II. THE REJECTION OF CLAIMS 1-7, 9, 11-16, 45, 48, 51-53, 56-59, 61-63 AND 65-78 UNDER 35 U.S.C. §112, FIRST PARAGRAPH – NEW MATTER**

Claims 1-7, 9, 11-16, 45, 48, 51-53, 56-59, 61-63 and 65- 78 are rejected under 35 U.S.C. §112, first paragraph, because it is alleged that the claimed subject matter is not described in the specification. Specifically, the Examiner states that in Claim 1, step (e) is not supported by the as-filed specification. Pursuant to MPEP 2163.06, the Examiner requests that applicants point out where in the original disclosure the amendments to the claims appear.

It appears that the same rejection that applies to claim 1, step (e) also applies to the other independent claims 53, 59 and 63. In particular, in each of these claims, the amendment that was made to step e) is underlined in the portion of step e) reproduced below, which recites:

(e) testing the identified protease(s) or biologically active portion thereof for cleavage and inactivation of an activity of the target protein that contains the substrate sequence, thereby identifying a protease mutein or a catalytically/biologically active portion thereof that inactivates an activity of a target protein that is involved with the disease or pathology.

Hence, the amended portion of the claims is directed to testing the activity of the target protein upon cleavage by an identified proteases to verify that the identified mutein protease inactivates an activity of the target protein. Basis for this amendment can be found at paragraph [0050] and [0059], which describe general methods of engineering a protease to inactivate an activity of a target protein, and paragraph [0123], which describes testing or assaying the target protein upon cleavage by the protease for inactivation or destruction of a function or activity.

[0050] The present invention is drawn to methods for generating and screening proteases to cleave target proteins at a given substrate sequence. Proteases are protein-degrading enzymes that recognize an amino acid or an amino acid substrate sequence within a target protein. Upon recognition of the substrate sequence, proteases catalyze the hydrolysis or cleavage of a peptide bond within a target protein. **Such hydrolysis of the target protein may inactivate it, depending on the location of peptide bond within the context of the full-length sequence of the target sequence.** The specificity of proteases can be altered through protein engineering. If a protease is engineered to recognize a substrate sequence within a target protein or proteins that would (i.) **alter the function *i.e.* by inactivation of the target protein(s) upon catalysis of peptide bond hydrolysis and, (ii.) the target protein(s) are recognized or unrecognized as points of molecular intervention for a particular disease or diseases, then the engineered protease has a therapeutic effect via a proteolysis-mediated inactivation event.** [emphasis added]

[0059] In some examples, the engineered protease is designed to cleave any of the target proteins in Table 1, **thereby inactivating the activity of the protein.** The protease can be used to treat a pathology associated with that protein, by inactivating one of the target proteins. [emphasis added]

[0123] Mutant proteases that match the desired specificity profiles, as determined by substrate libraries, are then assayed using individual peptide substrates corresponding to the desired cleavage sequence. Variant proteases are also assayed to ascertain that they will cleave the desired sequence when presented in the context of the full-length protein. **The activity of the target protein is also assayed to verify that its function has been destroyed by the cleavage event.** The cleavage event is monitored by SDS-PAGE after incubating the purified full-length protein with the variant protease. [emphasis added]

Therefore, the specification as originally filed describes a method for generating and screening proteases as claimed that includes a step of testing the identified mutant proteases for activity in cleaving the target protein.

### **III. THE REJECTION OF CLAIMS 1-7, 9, 11-16, 45, 48, 51-53, 56-59, 61-63 AND 65-78 UNDER 35 U.S.C. §112, FIRST PARAGRAPH -Lack of Support**

Claims 1-7, 9, 11-16, 45, 48, 51-53, 56-59, 61-63 and 65- 78 are rejected under 35 U.S.C. §112, first paragraph, because 1) the Examiner states that the specification fails to provide written description for candidate proteases that inactivate an activity of the target



protein, the inactivation of which can ameliorate a disease or pathology and 2) the Examiner also states the specification fails to describe method step (e). With respect to this latter rejection, the Examiner states:

Furthermore, the specification fails to describe the claim method of step (e) testing the identified protease(s) or biologically active portions thereof for cleavage and inactivation of an activity of the target protein that contains the substrate sequence, thereby identifying a protease mutein or a biologically active portion thereof that inactivates an activity of a target protein that is involved with the disease or pathology. It is not apparent from the disclosure description as to type or kind of testing of any identified protease, if any, that has been applied to any or all kinds of target protein involved in any type of disease or pathology.

This rejection respectfully is traversed.

### **Relevant Law**

The case law related to the written description requirement is set forth in previous responses, incorporated by reference herein.

### **The Rejected Claims**

There are four independent claims, claims 1, 53, 59 and 63, that differ with respect to the particular target proteins or scaffold proteases used in the method. The basic steps of the method are the same for all independent claims. For example, independent claim 1, as amended herein, recites:

A method of producing and identifying a mammalian protease mutein that inactivates an activity of a target protein involved with a disease or pathology in a mammal, comprising the steps of:

(a) producing a library comprising protease muteins comprising mutation(s) of a serine protease scaffold and/or biologically active portions thereof at an amino acid position selected from among amino acid residues between 58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering, wherein:

each different mutein protease in the library is a member of the library;

each member of the library has N mutations relative to a wild-type mammalian protease scaffold or a biologically active portion thereof; and

N is a positive integer;

(b) contacting members of the library with a target protein or with a polypeptide comprising a substrate sequence that is present in the target protein, wherein:

the target protein is selected from among a cell surface molecule that transmits an extracellular signal for cell proliferation, a cytokine, a cytokine receptor and a signaling protein that regulates apoptosis; and

cleavage of a substrate sequence in the target protein inactivates an activity of the target protein; and

(c) measuring a cleavage activity and/or substrate specificity of at least two members of the library for the target protein or substrate sequence;

(d) based on the measured activity and/or specificity, identifying members of the library that have an increased cleavage activity and/or altered substrate specificity for cleaving

the substrate sequence or the target protein, relative to the wild-type mammalian protease scaffold; and

(e) testing the identified protease(s) or biologically active portion thereof for cleavage and inactivation of an activity of the target protein that contains the substrate sequence, thereby identifying a protease mutein or a biologically active portion thereof that inactivates an activity of a target protein that is involved with the disease or pathology.

Dependent claims further recite particulars of the number of mutations, the target protein, the scaffold protease, and steps of measuring cleavage activity and inactivation of an activity of the target protein.

### **Analysis**

First, for clarity of the claims, the language “inactivation of the target protein can ameliorate a disease or pathology” is deleted from the claims. As stated in previous responses, the instant claims are directed to a **method of screening for mutant proteases that inactivate an activity of a target protein**. While the specification renders it clear that the identified proteases can serve as candidate therapeutics for treating a disease or pathology (see *e.g.* at paragraph [0005], which states that “the resultant proteins may be used as agents for in vivo therapy”), the instant claims are not directed to a method of treatment of diseases or pathologies, nor to methods of ameliorating a disease or pathology. Hence, the fact that the inactivation of the target protein can ameliorate a disease or pathology is not required by the instant claims, and is not a limitation of the method of the instant claims. Applicant’s inclusion of this limitation was only meant to render it clear that the target protein is involved in a disease or pathology by virtue of the fact that destruction or inactivation of an activity of the target protein can ameliorate a disease or pathology. Accordingly, by deletion of this phrase, the portion of the rejection that relates to written description for candidate proteases that inactivate an activity of the target protein, the inactivation of which can ameliorate a disease or pathology, is rendered moot.

Second, with respect to the rejection as to the description of step e) of the claims, Applicant respectfully submits that the specification provides detailed description, including in Working Examples, sufficient to show that Applicant has possession of the claimed subject matter. In particular, the specification describes exemplary target proteins for use in the method and also describes assays for testing the activity of a target protein in order to assay whether a function or activity of the target protein has been destroyed or inactivated following cleavage by protease. Further, as set forth in MPEP 2163, Applicant is not required to describe information that is well known in the art (stating that “[i]nformation which is well known in the art need not be described in detail in the specification. See, *e.g.*,

*Hybritech, Inc. v. Monoclonal Antibodies, Inc.*, 802 F.2d 1367, 1379-80, 231 USPQ 81, 90 (Fed. Cir. 1986)). In this instance, as described further below, assays to assess or test the function or activity of a target protein is well known in the art.

The specification describes that exemplary of target proteins are cell surface molecules that transmit an extracellular signal, cytokines, cytokine receptors, and proteins involved in apoptosis. For example, a cell surface molecule that transmits an extracellular signal is a G protein coupled receptor (GPCR) or a kinase. One of skill in the art is familiar with target proteins that are cell surface molecules that transmit an extracellular signal, cytokines, cytokine receptors, and proteins involved in apoptosis. The specification describes at least 45 exemplary target proteins that are representative of these classes and that are involved with a disease or pathology (*see e.g.* Table 1 on page 9). Table 1 is reproduced below:

**Table 1**

Target	Indication	Molecule class
IL-5/IL-5R	Asthma	Cytokine
IL-1/IL-1R	Asthma, inflammation, Rheumatic disorders	Cytokine
IL-13/IL-13R	Asthma	Cytokine
IL-12/IL-12R	Immunological disorders	Cytokine
IL-4/IL-4R	Asthma	Cytokine
TNF/TNFR	Asthma, Crohn's disease, HIV infection, inflammation, psoriasis, rheumatoid arthritis	Cytokine
CCR5/CXCR4	HIV infection	GPCR
gp120/gp41	HIV infection	Fusion protein
CD4	HIV infection	Receptor
Hemagglutinin	Influenza infection	Fusion protein
RSV fusion protein	RSV infection	Fusion protein
B7/CD28	Graft-v.-host disorder, rheumatoid arthritis, transplant rejection, diabetes mellitus	Receptor
IgE/IgER	Graft-v.-host disorder, transplant rejection	Receptor
CD2,CD3,CD4,CD40	Graft-v.-host disorder, transplant rejection, psoriasis,	Receptor
IL-2/IL-2R	Autoimmune disorders, graft-v.-host disorder, rheumatoid arthritis	Cytokine
VEGF,FGF,EGF,TGF	Cancer	Cytokine
Her2/neu	Cancer	Receptor
CCR1	Multiple sclerosis	GPCR
CXCR3	Multiple sclerosis, rheumatoid arthritis	GPCR
CCR2	Atherosclerosis, rheumatoid arthritis	GPCR
Src	Cancer, osteoporosis	Kinase
Akt	Cancer	Kinase
Bcl-2	Cancer	Protein-protein
BCR-Abl	Cancer	Kinase
GSK-3	Diabetes	Kinase
cdk-2/cdk-4	Cancer	Kinase

Other exemplary proteins involved in apoptosis also are provided in the specification at page 39, lines 19-24, i.e. human cytochrome c, human Apaf-1, human caspase-9, human caspase-7, human caspase-6, human caspase-2, human BAD, human BID, human BAX, human PARP, or human p53. These target proteins are not new, but are proteins that are well known in the art and have activities that are well known in the art. **The instant claims explicitly recite these classes of target proteins or these target proteins.** Further, the description of these target proteins is representative of a class of target proteins, such that one of skill in the art would understand and be able to identify a target protein involved in a pathology or disease.

In this case the modified proteases are screened for cleavage and inactivation of an activity of the target protein. The specification describes that in the method, proteases are engineered to alter the function of a target protein to inactivate it (see *e.g.* at page 9, lines 16-19). Thus, the specification describes that in the method “[t]he activity of the target protein is also assayed to verify that its function has been destroyed by the cleavage event” (see *e.g.* at page 32, lines 23-25).

The specification exemplifies activities or functions of exemplary proteins that can be inactivated or destroyed upon incubation and cleavage by a protease. For example, the specification describes that cleavage and inactivation of a TNF receptor can inactivate the receptor such that the action of tumor necrosis factor is inhibited (see *e.g.* at page 9, lines 29-32). In another example, the specification describes that cleavage and inactivation of a target that is the same as activated protein C, which is involved in promoting blood coagulation, can attenuate the blood coagulation cascade (see *e.g.* at page 10, lines 1-3). In a further example, the specification describes that cleavage and inactivation of a cell surface molecule involved in tumorigenicity can inactivate their ability to transmit extracellular signals, especially for cell proliferation (see *e.g.* at page 10, lines 4-6). In an additional example, the specification describes that cleavage and inactivation of a protein involved in cell cycle progression, can inactivate the ability of the protein to allow the cell cycle to go forward (see *e.g.* at page 10, lines 9-12). In other examples, the specification describes that cleavage and inactivation of viral membrane fusion proteins can inhibit the ability of the virus to infect cells (see *e.g.* at page 10, lines 13-17). In further examples, the specification describes that cleavage and inactivation of the same target protein as plasminogen activator can result in inactivation of the thrombolytic cascade (see *e.g.* at page 10, lines 18-21). In additional examples, the specification describes that cleavage and inactivation of cytokines or receptors can inactivate the associated signaling cascade (see *e.g.* at page 10, lines 22-26). In other examples, the



specification describes that cleavage and inactivation of proteins involved in apoptosis inhibit the associated signaling cascade for apoptosis and inhibit apoptosis (*see e.g.* at page 10, lines 17-31 and at page 39, lines 18-24).

Accordingly, there is extensive descriptions of functions that can be altered by inactivation of a target protein. It is within the level of one of skill in the art to assay for such functions depending on the specific target protein. For example, as described in the specification, following incubation with a protease, various functions of a target protein including, for example, cell proliferation, coagulant activity, binding of a ligand to a receptor (*e.g.* TNF binding to TNF receptor), cell proliferation, other assays for cell cycle activity, viral infection of cells, and apoptosis can be assayed. As set forth below, *in vitro* and *in vivo* assays to assess various functions of target proteins, including target proteins as exemplified in the instant specification, were well known to one of skill in the art at the time of the priority date of the instant application:

Target	Assay	Reference
Caspase 3	ApoAlert CPP32/Caspase-3 colorimetric assay kit (Clontech)	Izban et al., "Characterization of the interleukin-1 beta-converting enzyme/ced-3-family protease, caspase-3/CPP32, in Hodgkin's disease: lack of caspase-3 expression in nodular lymphocyte predominance Hodgkin's disease," Am J Path 154(5):1439-1447 (1999)
tumor necrosis factor	<i>In vitro</i> TNF assay with TNF-sensitive L cells; <i>In vivo</i> TNF assay – standard Meth A sarcoma assay.	Rubin et al., "Purification and characterization of a human tumor necrosis factor from the LuKII cell line," Proc Natl Acad Sci USA 82:6637-6641 (1985); Williamson et al., Proc Natl Acad Sci USA 80:5397-5401 (1983)
interleukin-1	Thymocyte proliferation assay; Ability to stimulate IL-2 production	Lachman et al., J. Immunol. 119:2019 (1977); Howard et al., "Role of interleukin 1 in anti-immunoglobulin-induced B cell proliferation," J. Exp Med 157:1529-1543 (1983); Lowenthal and MacDonald, "Binding and internalization of interleukin 1 by T cells," J. Exp Med 164:1060-1074 (1986)
interleukin-2	IL-2 dependent exponential proliferation of murine cytotoxic T-cell line (CTLL)	Gillis and Mizel, "T-Cell lymphoma model for the analysis of interleukin 1-mediated T-cell activation," Proc Natl Acad Sci USA 78(2):1133-1137 (1981)
interleukin-4	Cellular Assay – stimulate proliferation of purified B cells	Grabstein et al., "Purification to homogeneity of B cell stimulating factor. A molecule that stimulates proliferation of multiple lymphokine-dependent cell lines," J Exp Med, 163:1405-1414 (1986)

interleukin-4 receptor	Induction of intracellular signaling	Kammer et al., "Homodimerization of interleukin-4 receptor $\alpha$ chain can induce intracellular signaling," J Biol Chem 271(39):23634-23637 (1996)
interleukin-5	cytotoxicity in LAK cells	Aoki et al., "Interleukin 5 enhances interleukin 2-mediated lymphokine-activated killer activity," J Exp Med, 170:583-588 (1989)
interleukin-12	chemotactic activity	Allavena et al., "Interleukin-12 is chemotactic for natural killer cells and stimulates their interaction with vascular endothelium," Blood 84:2261-2268 (1994)
interleukin-13	IL-13 induced IL-6 secretion in Caki-1 cells	Caput et al., "Cloning and Characterization of a Specific Interleukin (IL)-13 Binding Protein Structurally Related to the IL-5 Receptor $\alpha$ Chain," J Biol Chem 271(28): 16921-16926 (1996)
p-selectin	Induction of tissue factor expression	Celi et al., "P-selectin induces the expression of tissue factor on monocytes," Proc Natl Acad Sci USA 91:8767-8771 (1994)
p-selectin glycoprotein ligand (PSGL)	Binding to immobilized p-selectin	Moore et al., "The P-selectin Glycoprotein Ligand from Human Neutrophils Displays Sialylated, Fucosylated, O -Linked Poly-N-acetyllactosamine," J Biol Chem 269:22318-23327 (1994)
Substance P	Substance P with human blood T-lymphocytes stimulates T-lymphocyte proliferation, quantified by flow cytometric and direct binding assays.	Payan et al., "Substance P recognition by a subset of human T lymphocytes," J Clin Invest. 74:1532-1539 (1984)
bradykinin	Effect on rat dorsal root ganglion neurons in vitro	Thayer et al., "Regulation of calcium homeostasis in sensory neurons by bradykinin," J Neuroscience 8(11):4089-4097 (1988)
Factor IX	coagulation assay	Chung et al., "purification and characterization of an abnormal Factor IX (Christmas factor) molecule," J Clin Invest 62:1078-1085 (1978)
CCR5	Fusion assay in HeLa-P4 cells	Picard et al., "Multiple extracellular domains of CCR-5 contribute to human immunodeficiency virus type 1 entry and fusion," J Virology 71:5003-5011 (1997)
CXCR4	Functional assay in U373MG-CD4 cells	Brelot et al., "Role of the first and third extracellular domains of CXCR-4 in human immunodeficiency virus coreceptor activity," J Virology 71:4744-4751 (1997)
gp120	CD4 binding	Willey et al., "In vitro mutagenesis identifies a region within the envelope gene of the human immunodeficiency virus that is critical for

		infectivity," J Virology 62:139-147 (1988)
gp41	HIV replication in Jurkat lymphocytes; efficiency of viral entry	Gabuzda et al., "Effects of deletions in the cytoplasmic domain on biological functions of human immunodeficiency virus type 1 envelope glycoproteins," J Virology 66:3306-3315 (1992)
hemagglutinin	Agglutination	Sekikawa and Lai, "Defects in functional expression of an influenza virus hemagglutinin lacking the signal peptide sequences," Proc Natl Acad Sci USA 80:3563-3567 (1983)
B7	CD2- and CD3-dependent T-cell proliferation	Azuma et al., "CD28 interaction with B7 costimulates primary allogeneic proliferative responses and cytotoxicity mediated by small, resting T lymphocytes," J Exp Med 175:353-360 (1992)
CD28	PI3-kinase activity	Pages et al., "Two distinct intracytoplasmic regions of the T-cell adhesion molecule CD28 participate in phosphatidylinositol 3-kinase association," J Biol Chem 271:9403-9409 (1996)
CD2	IL-2 production	Chang et al., "Dissection of the human CD2 intracellular domain. Identification of a segment required for signal transduction and interleukin 2 production," J Exp Med 169:2073-2083 (1989)
CD4	Inhibition of syncytium formation and virus replication	Thali et al., "Effects of changes in gp120-CD4 binding affinity on human immunodeficiency virus type 1 envelope glycoprotein function and soluble CD4 sensitivity," J Virology 65:5007-5012 (1991)
VEGF	Induction of expression of connective tissue growth factor	Suzuma et al., "Vascular endothelial growth factor induces expression of connective tissue growth factor via KDR, Flt1, and phosphatidylinositol 3-kinase-akt-dependent pathways in retinal vascular cells," J Biol Chem 275:40725-70731 (2000)
fibroblast growth factor	Proliferation of capillary endothelial cells	Feige and Baird "Basic fibroblast growth factor is a substrate for protein phosphorylation and is phosphorylated by capillary endothelial cells in culture," Proc Natl Acad Sci USA 86:3174-3178 (1989)
EGF receptor	Cell proliferation	Moriyai et al., "A variant epidermal growth factor receptor exhibits altered type alpha transforming growth factor binding and transmembrane signaling," Proc Natl Acad Sci USA 91:10217-10221 (1994)
transforming growth factor	Inhibition of TNF- $\alpha$ production	Ranges et al., "Inhibition of cytotoxic T cell development by transforming growth factor beta and reversal by recombinant tumor necrosis factor alpha," J Exp Med 166:991-998 (1987)

Hence, based on the description in the specification, it is within the level of one of skill in the art to test for inactivation of an activity of a target protein. As an example, the

specification describes numerous assays that can be used to assess apoptosis including, for example, DNA ladder formation by gel electrophoresis, morphologic examination by electron microscopy, flow cytometry including with the use of various DNA dyes, Annexin V or TUNNEL staining (see *e.g.* at page 39, lines 3-17).

Further, Working Examples are provided that evidence assays to assess inactivation of an activity of a target protein following incubation with a protease. For example, Example 4 is directed to proteolytic cleavage and inactivation of tumor necrosis factor and tumor necrosis factor receptor (TNFR) and describes various assays to assess cleavage and inactivation of the cytokine receptor or cytokine. For example, the Example describes assessing inactivation of TNFR by cleavage of its stalk region such that it becomes free from the membrane and released into the media; the released protein can be detected in the media by ELISA. Also, the Example describes an assay to assess cleavage and inactivation of TNF by looking for degradation products of a radioactive protein. The Example also describes assays of further testing for cleavage and inactivation of the ligand or receptor by incubation of the radioactive cytokine ligand or neutrophils (PMNs; expressing the TNFR) with protease. Following incubation with protease the radioactive ligand is incubated with the neutrophils, and binding to the TNFR thereon is determined.

Example 8 further describes various assays to screen for cleavage of target proteins and inactivation or destruction of a function of the target protein. As described in the Example, following incubation of a purified target protein or cells expressing the target protein, cleavage can be assessed by assays such as SDS-PAGE and visualization of protein by Coomassie blue staining, autoradiography using a radiolabeled probe or by Western blot. ELISA also can be used. The Example further exemplifies this with respect to cleavage of tumor necrosis factors, TNFR1 and TNFR2.

Example 9 further describes functions associated with various cytokines and cytokine receptors, which can be assayed for to determine if cleavage by a protease destroyed the function. For example, Example 9 describes that TNF its receptor TNFR (TNFR1 or TNFR2) are associated with various signal transduction events such as the production of pro-inflammatory cytokines (IL-1, IL-6 and GM-CSF), induction of metalloproteinases such as collagenase and stromelysin, and increased in proliferation and activity of osteoclasts. Any of these functions can be assessed to determine if they are destroyed or inactivated upon incubation of the cytokine or receptor with protease. The Example further describes that VEGF/VEGFR is associated with angiogenesis and that the receptors are tyrosine kinases



involving in initiating signaling cascades. Hence, such functions can be assessed to determine if they are destroyed or inactivated upon incubation of a VEGF or receptor with protease. The Example further describes functions and activities of the ligands EGFR and HER-2 and other related ErbB family of receptor tyrosinse kinases. The Example describes that activation of the ErbB receptors through binding of ligand results in downstream signaling through the mitogen-activated protein kinase (MAPK) and the AKT/phosphoinositide 3-kinase (PI3-kinase) pathways. Also, the Example describes that these receptors are involved in cell proliferation, differentiation and angiogenesis. Hence, any of these functions can be assessed to determine if they are destroyed or inactivated upon incubation of the ligand or receptor with protease.

Finally, Example 11 exemplifies cleavage and inactivation of an activity of the target protein caspase-3 by an exemplified granzyme B mutant (I99A/N218A). It is noted that granzyme B normally cleaves caspase-3 to activate it. This Example, however, describes cleavage of a distinct substrate sequence in caspase-3 by the mutant protease such that activity of caspase-3 is inactivated or destroyed (see *e.g.* at page 52 lines 1-5 and Figure 1, which demonstrate the altered specificity of the mutant protease). This inactivation cleavage sequence is only recognized by the mutant granzyme B and not by wildtype granzyme B (see *e.g.* at page 52, lines 22-27). In particular, the Example and accompanying Figure 7B exemplify that granzyme B cleaved and inactivated full-length caspase-3 as assessed in assays for apoptosis activity. For example, the Example states “[a]s shown in Figure 7B, the mutant granzyme B antagonized the effect of wildtype granzyme B to induce apoptosis by inactivating caspase-3.”

In sum, it respectfully is submitted that the specification renders it clear that the method of identifying a mutant protease that inactivates a target protein, includes a step of testing or assaying a function or activity of the target protein to verify that it has been destroyed following incubation by the protease. The specification provides numerous examples of exemplary target proteins that are associated with a disease or pathology, and in particular numerous examples that are representative of the class of target proteins that provide an extracellular signal, a cytokine, a cytokine receptor or a protein involved in apoptosis. The specification describes the function of representative of such target proteins, and also describes the consequences of inactivating an activity of the protein. The specification, including in Working Examples, describes several assays to assess the function of target proteins to determine if an activity has been destroyed or inactivated, such as cell

signaling assays, binding assays, proliferation assays and apoptosis assays. Further, functions of target proteins, including the exemplary classes of target proteins described in the specification, and assays to assess the function of the target proteins are well known to one of skill in the art. Accordingly, it respectfully is submitted that Applicant has possession of a method that includes as a step "testing the identified protease(s) or biologically active portion thereof for cleavage and inactivation of an activity of the target protein."

**IV. THE REJECTION OF CLAIMS 1-7, 9, 11-16, 45, 48, 51-53, 56-59, 61-63 AND 65-78 UNDER 35 U.S.C. §112, FIRST PARAGRAPH - Scope**

Claims 1-7, 9, 11-16, 45, 48, 51-53, 56-59, 61-63 and 65-78 are rejected under 35 U.S.C. §112, first paragraph, as lacking written description because the Examiner states that the specification describes granzyme as the sole cleaving enzyme used in the method, but that the claims encompass an enormous variation of enzyme muteins, even for granzyme, which contains numerous allelic variants. The Examiner states that the genus of proteases encompassed by the claims as employed in the method is too large and structurally diverse. Hence, the Examiner asserts that:

[T]he instant claims are not adequately described in the specification to claim the broad and structurally diverse genus of the compounds used in the method. The claims lack written description and a skilled artisan cannot, as one can do with a fully described genus, visualize or recognize the identity of the members of the genus that are allelic variants. This is true even for the single granzyme, let alone for any enzymes of no claim distinguishing features as use in the broad methods steps.

Applicant respectfully traverses this rejection. The arguments and response filed August 7, 2008, June 11, 2009 and July 20, 2010 are incorporated by reference.

**Relevant Law**

The case law related to the written description requirement is set forth in previous responses, incorporated by reference herein.

**The Rejected Claims**

Claim 1 is set forth above. Independent claims 53, 59 and 63 also are directed to a **method of screening** to produce and identify a protease that inactivates an activity of a target protein involved in a disease or pathology, but differ in their scope. Each of the claims include steps of a) producing a library of protease muteins; b) contacting mutein protease members of the library with a substrate sequence in a target protein or a target protein involved in a disease or pathology; c) measuring a cleavage activity and/or substrate specificity of at least two members; d) based on the measured activity and/or specificity

identifying members that have increased cleavage activity or specificity compared to the starting protease scaffold; and e) testing the identified proteases or biologically active portion thereof for cleavage and inactivation of an activity of the target protein, thereby identifying a mutein protease that cleaves and inactivates an activity of a target protein involved in a disease or pathology.

As amended, **all claims** recite that in the step of producing a library of protease mutein of protease scaffolds, the protease muteins include mutation of a serine protease or a particular recited serine protease **at an amino acid position selected from among amino acid residue between 58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering**. Hence, all claims provide the structure of the protease muteins included in the library that is used in the method.

The independent claims 1, 53, 59 and 63 differ in their recitation of the particular protease scaffold used in the method, or in the target protein that is a target of the method. For example, independent **claim 1** is directed to **specific target proteins** such as a cell surface molecule that transmits an extracellular signal for cell proliferation, a cytokine, a cytokine receptor and a signaling protein that regulates apoptosis and also recites that the target protein is a **serine protease scaffold**. Dependent claim 7 further recites specific serine protease scaffolds. Independent **claim 53** is directed to **specific mammalian protease scaffolds** such as a granzyme A, granzyme B, granzyme M, cathepsin, trypsin, chymotrypsin, subtilisin, MTSP-1, elastase, chymase, tryptase, collagenase, papain, neutrophil elastase, complement factor serine proteases, ADAMTS13, neural endopeptidase/neprilysin, furin, cruzain, and urokinase plasminogen activator (uPA), and also recites further steps of the method f) –i) **Independent claim 59** is directed to **specific target proteins** such as a cell surface molecule that transmits an extracellular signal for cell proliferation, a cytokine, a cytokine receptor and a signaling protein that regulates apoptosis and also directed to specific protease scaffolds such as granzyme A, granzyme B, granzyme M, cathepsin, trypsin, chymotrypsin, subtilisin, MTSP-1, elastase, chymase, tryptase, collagenase, papain, neutrophil elastase, complement factor serine proteases, ADAMTS13, neural endopeptidase/neprilysin, furin, cruzain, and urokinase plasminogen activator (uPA). **Independent claim 63** is directed to **specific recited target proteins** and also is directed to **specific human protease scaffolds**.

#### **Analysis**

As discussed in the Guidelines for Examination of Patent Applications Under the 35

U.S.C. 112, para. 1, "Written Description" Requirement (hereinafter the Guidelines), the written description requirement for a claimed genus may be satisfied through 1) sufficient description of a representative number of species by actual reduction to practice, 2) reduction to drawings, or 3) by disclosure of relevant, identifying characteristics, *i.e.*, structure or other physical and/or chemical properties, by functional characteristics coupled with a known or disclosed correlation between function and structure, or by a combination of such identifying characteristics, sufficient to show the applicant was in possession of the claimed genus.

In this instance, Applicant has more than complied with the written description requirement by providing sufficient description of a representative number of species **and** by providing relevant identifying characteristics of all members of the genus.

With respect to the protease scaffolds to which the protease muteins contain mutations of, it respectfully is submitted that the instant claims are directed to methods of screening libraries that contain muteins of proteases. Applicant is not claiming the protease scaffolds or the mutant proteases, but methods in which proteases are used as scaffolds to create mixtures. The skilled artisan is capable of preparing mixtures of mutant proteases (or of catalytically active portions thereof). Screening methods and libraries by their nature rely on diversity, not knowledge of any structure/function relationships; the screening methods ferret out products that have a desired activity. Notwithstanding this, the specification teaches and exemplifies a wide variety of protease scaffolds and also teaches residues that are specificity determinants into which mutations can be made, and how to make mutations. By virtue of mutation of amino acid residues that are associated with specificity of proteases, the structure of the protease muteins in the library is correlated with function.

For example, the specification describes exemplary scaffold proteases for use in the method (*see e.g.* Table 2). These are all known proteases in the art, including known allelic and species variants. **Applicant is not inventing these proteases; they are known.** As stated above and in previous responses, Applicant is not required to describe information that is well known in the art (*see, e.g., Hybritech, Inc. v. Monoclonal Antibodies, Inc.*, 802 F.2d 1367, 1379-80, 231 USPQ 81, 90 (Fed. Cir. 1986)). Also, there is no per se rule to provide sequence information when that sequence is already known in the field. Capon v. Eshhar v. Dudas 418 F.3d 1349 (Fed. Cir. 2005); Falkner v. Inglis 448 F.3d 1357, 1367-68 (Fed. Cir. 2006). As evidence of the knowledge of one of skill in the art with regard to the sequence and structure of various serine proteases, a list of references and accession numbers



evidencing the knowledge of one of skill in the art with respect to the sequences of known proteases is provided:

<b>PROTEASE (species)</b>	<b>Sequence Accession Number</b>
<b>Trypsin</b>	
Sus scrofa	P00761
Homo sapiens	P07477 Trypsin-1
Homo sapiens	P07478 Trypsin-2
Homo sapiens	P35030 Trypsin-3
Homo sapiens	AAA61231
Homo sapiens	AAG30943
Homo sapiens	AAC50728
Homo sapiens	AAA61232
Streptomyces griseus	P00775
Fusarium oxysporum	P35049
<b>Chymotrypsin</b>	
Homo sapiens	Q99895 Chymotrypsin-C
Homo sapiens	P17538 Chymotrypsinogen B
Homo sapiens	AAA52128 Chymotrypsinogen B
Homo sapiens	AAP36020 Chymotrypsinogen B
Rattus norvegicus	P07338 Chymotrypsin B
Gadus morhua (Atlantic cod)	P80646 Chymotrypsin B
Vespa orientalis (Oriental hornet)	P00768 Chymotrypsin-2
Vespa crabro (European hornet)	P00769 Chymotrypsin-2
Phaedon cochleariae (Mustard beetle)	O97398
Rattus norvegicus	P55091 Chymotrypsin C
<b>Subtilisin</b>	
Bacillus subtilis	P04189 Subtilisin E
Bacillus subtilis	AAA22742
Bacillus subtilis	AAA22814
Bacillus subtilis	AAA22744
Bacillus licheniformis	P00781 Subtilisin DY
Bacillus lentus	P29599 Subtilisin BL
<b>Granzyme A</b>	
Homo sapiens	P12544
Homo sapiens	AAA52647
Homo sapiens	AAD00009
Mus musculus	P11032
<b>Granzyme B</b>	
Homo sapiens	P10144
Homo sapiens	AAA52118
Homo sapiens	AAB59528
Homo sapiens	AAA36627
Homo sapiens	AAA67124
Mus musculus	P04187
<b>Granzyme M</b>	
Homo sapiens	P51124
Homo sapiens	AAA59582

Homo sapiens	AAA57262
Rattus norvegicus	Q03238
Mus musculus	O08643
<b>Elastase</b>	
Gadus morhua (Atlantic cod)	P32197
<b>Chymase</b>	
Homo sapiens	P23946
Homo sapiens	AAB26828
Homo sapiens	AAA52020
Homo sapiens	AAA52021
Homo sapiens	AAA52019
Canis familiaris	P21842
Mus musculus	P21844
Macaca fascicularis	P56435
Papio hamadryas	P52195
Rattus norvegicus	P50339
<b>Papain</b>	
Carica papaya	P00784
<b>Neutrophil elastase</b>	
Homo sapiens	P08246
Homo sapiens	AAA36359
Homo sapiens	AAA36173
Equus caballus	P37357 Neutrophil elastase 2A
Equus caballus	P37358 Neutrophil elastase 2B
Mus musculus	Q9Z284
<b>Complement factor serine proteases</b>	
Homo sapiens	P48740 MASP1
Mus musculus	P98064 MASP1
Homo sapiens	AAK84071
<b>ADAMTS13</b>	
Homo sapiens	Q96P40
Homo sapiens	BAB69487
<b>Neural endopeptidase/neprilysin</b>	
Homo sapiens	P08473
Homo sapiens	AAA51915
Homo sapiens	CAA68752
Homo sapiens	AAA52294
Homo sapiens	CAA30157
Mus musculus	Q61391
Oryctolagus cuniculus	P08049
Rattus norvegicus	P07861
Sus scrofa	P19621
<b>Furin</b>	
Homo sapiens	P09958
Mus musculus	P23188
Bos taurus	Q28193
Rattus norvegicus	P23377
Xenopus laevis	P29119

<b>Cruzain</b>	
Trypanosoma cruzi	P25779
Trypanosoma cruzi	AAA30269
Trypanosoma cruzi	AAA30270
Trypanosoma cruzi	CAA38278
Trypanosoma cruzi	AAA30180
<b>Urokinase plasminogen activator (uPA)</b>	
Homo sapiens	P00749
Homo sapiens	AAA61253
Homo sapiens	CAA26535
Homo sapiens	AAC97138
Mus musculus	P06869
Oryctolagus cuniculus	Q8MHY7
Rattus norvegicus	P29598
Bos Taurus	Q05589
Papio cynocephalus	P16227
Sus scrofa	P04185
Gallus gallus	P15120

**In addition, the specification and claims provide the relevant identifying characteristics of the protease muteins used in the method that correlate structure and function.** In particular, the specification describes that amino acid residues that contribute to the substrate specificity can be targeted for mutagenesis, *i.e.* those amino acids that are specificity determinants. These are amino acid residues in proteases that are identified in the application conferring substrate specificity to the protease; hence, mutating these residues can alter or change the substrate specificity. The specification describes exemplary specificity determinants for serine proteases (see *e.g.* at page 20, lines 3-7; and at page 28, lines 1-3; and Table 3). These specificity determinants include, for example, amino acid residues between 58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering. The chymotrypsin numbering scheme was devised by Jonathan Greer (PROTEINS: Structure, Function and Genetics, 7:317-334 (1990)) and is based on sequence alignment between and among various serine proteases with numbering based on the numbering scheme of bovine chymotrypsin A. Alignment of exemplary serine proteases, including trypsin, and chymotrypsin and identification of corresponding residues by chymotrypsin numbering is exemplified in see, *e.g.* Friedrich *et al.* J Biol Chem. 2002 Jan 18;277(3):2160-8. The claims, as amended, recite that the proteases contain mutations at amino acid residues between 58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering. The specification further describes

exemplary granzyme B mutants containing mutations in these specificity determinants (see *e.g.* Table 4).

Under the heading, "Mutagenesis of the Scaffold Protease," the specification describes that mutation of specificity determinants can be effected as follows (see *e.g.*, page 25, lines 3-16):

In order to change the substrate preference of a given subsite (S1-S4) for a given amino acid, the specificity determinants that line the binding pocket are mutated, either individually or in combination. In one embodiment of the invention, a saturation mutagenesis technique is used which the residue(s) lining the pocket is mutated to each of the 20 possible amino acids... Alternatively, single amino acid changes are made using standard, commercially available site-directed mutagenesis kits such as QuikChange (Stratagene). In another embodiment, any method commonly known in the art for specific amino acid mutation could be used...

For example, the specification describes that each member of the library is a protease scaffold with a number (N) of mutations made to each member of the library. The specification describes that the resulting members in the library can contain a number of mutations such as, for example, 1, 2-5 (*e.g.* 2, 3, 4 or 5), 5-10 (*e.g.* 5, 6, 7, 8, 9 or 10), or 10-20 (*e.g.* 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20) (see *e.g.* at page 3, lines 709).

Hence, the specification describes the scaffold proteases (which are known proteases well known to one of skill in the art), describes residues thereof that are specificity determinants and that are targeted for mutagenesis in the protease muteins, describes methods of effecting mutations, and exemplifies mutation of specificity determinants in granzyme B. Hence, one of skill in the art can easily envision the full scope of the claimed subject matter by knowing the sequence of the protease scaffold and the mutant residues contained in the library of protease muteins. Accordingly, it respectfully is submitted that because Applicant adequately describes the scaffold to be mutagenized, how to effect mutations in order to produce a library of protease muteins and what residues to mutate, Applicant had possession of the claimed methods, which are for identifying mutein proteases, at the time of filing the application.

#### **Rebuttal to Examiner's Comments**

*1) The Examiner's Statement that "it is known that in biotechnology art one species cannot be extrapolated to another due to the numerous unforeseeable reaction or interactions or protease-substrate" is misplaced*



The Examiner appears to be stating that the genus of proteases is not described because there is not any correlation of function between the proteases in the method. For example, the Examiner states as an example citing Rosen *et al.* (20020086811) that:

caspases are among the most specific of proteases, with an unusual and absolute requirement for cleavage after aspartic acid. (The only other eukaryotic protease known to have a similar specificity is the serine protease granzyme B, a mediator of granule-dependent cytotoxic T lymphocyte-mediated apoptosis). Recognition of at least four amino acids NH<sub>2</sub>-terminal to the cleavage site is also a necessary requirement for efficient catalysis. The preferred tetrapeptide recognitions motif differs significantly among caspases and explains the diversity of their biological functions. Their specificity is even more stringent: not all proteins that contain the optimal tetrapeptide sequence are cleaved, implying that tertiary structural elements may influence substrate recognition. Cleavage of proteins by caspases is not only specific, but also highly efficient. The strict specificity of caspases is consistent with the observation that apoptosis is not accompanied by indiscriminate protein digestion; rather, a select set of proteins is cleaved in a coordinated manner, usually at a single site, resulting in a loss or change in function.

The Examiner is reminded that the claims are directed to a method of screening by engineering proteases to alter or change their substrate specificity. Hence, the proteases muteins identified by practice of the method have **new specificities and activities**, and hence extrapolation to any known or existing protease misses the point of the method. The method is a method to identify proteases with altered substrate specificity such that the protease cleaves and inactivates a target protein involved in a disease or pathology. By virtue of the change in specificity compared to the scaffold protease, the protease mutein can be used as a candidate therapeutic to destroy or inactivate the activity of the target protein. The substrate specificity of the scaffold protease is irrelevant to the claimed method, since the method is a method to identify protease muteins with new and altered substrate specificities. Hence, knowledge of the substrate specificity of the protease scaffold used in the method is not required. As discussed further below in rebuttal 2) there is no requirement in the claims that the protease scaffold have any cleavage activity and/or specificity for the target protein. Further, as stated above, the claims provide the relevant identifying characteristics of the mutein proteases by identifying residues that are specificity determinants for mutations; mutation of a residue associated with the specificity of a protease is necessarily correlated to the function of protease specificity, which is the function that is being engineered by practice of the method.

2) *The Examiner's Statement on page 15 of the Response that "[t]he mutated protease has to be screened against a target protein specific only for said protease in order to identify mutant(s) with increased activity relative to the parent protease" is not correct*

The Examiner appears to believe that the claimed method is limited to assessing cleavage activity and specificity of a protease mutein against a known target substrate of the corresponding protease scaffold. **This is not correct.** There is no limitation in the claims that recites that the target protein involved in a disease or pathology is a target protein of the protease scaffold. The point of the method is to engineer new and different specificities for protease mutants. Indeed, cleavage activity and/or substrate specificity for a target protein can be "increased" whether or not the corresponding protease scaffold exhibits any cleavage activity and/or substrate specificity for the target protein. For example, if the corresponding unmodified protease scaffold **does not** exhibit any cleavage activity and/or substrate specificity for a substrate sequence in a target protein or a target protein, then necessarily any protease mutein that **does** exhibit cleavage activity and/or substrate specificity for a substrate sequence in a target protein or a target protein exhibits "increased" cleavage activity and/or substrate specificity compared to the protease scaffold. Hence, there absolutely is no requirement in the claims that the protease scaffold exhibit any activity towards the target protein used in the method.

Accordingly, as stated in previous responses, the structure/function of the starting protease is not relevant to practice of the method; the method identifies those proteases with the desired function, *i.e.* increased cleavage and/or substrate specificity for a target protein compared to the corresponding scaffold protease. Yet, the structure/function of the protease muteins is clearly set forth in the specification, and in the claims as amended, by reciting that the mutations in the protease occur at residues that are specificity determinants, *i.e.* between residue 58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering. Hence, mutating the structure of the protease at these known identified residues is correlated to the function of altered substrate specificity and/or cleavage activity.

3) *Reliance on In re Ruschig, 379 F. 2d 990, 995, 154 USPQ 118, 23 (CCPA 1967) is inapt.*

The Examiner rebuts Applicant's arguments that the specification describes in detail the protease scaffold for use in the method (*e.g.* as set forth in Table 2), residues that can be mutated in generating the protease mutein libraries and methods of mutagenesis by stating that:

...general statements in the specification and/or definitions do not replace the requirement of the law that applicants describe in detail that which is claimed to show possession of the claims being sought. Furthermore, a listing or selection of every possible protease scaffold or target does not constitute a written description of every species in a genus. It would not “reasonably lead” those skilled in the art to any particular species. In *re* Ruschig, 379 F. 2d 990, 995, 154 USPQ 118, 23 (CCPA 1967).

The Examiner’s attention is directed to pages 25-26 of the response filed August 07, 2008 in which Applicant rebutted this exact same point. In that response, Applicant stated that the issue of the Court in *In re Ruschig* was whether the compound of claim 13 was supported by the disclosure, whether it was new matter. The disclosure did not disclose the specific compound of claim 13, but instead disclosed a general formula of a compound with specified R groups. It was argued that one of skill in the art could arrive at the claimed compound from the general formula. The Court, however, affirmed the Board of Appeals ruling that the compound was not supported.

The claims and disclosure at issue are not analogous to *In re Ruschig*. In the instant case, Applicant is not claiming a product by a general formula of proteases. Rather, Applicant is claiming a **method** that can employ any protease as a scaffold, regardless of its function or substrate specificity, in order to generate a library of protease muteins that contain mutations at amino acid residues between 58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering. The specification provides an extensive list of exemplary scaffold proteases that can be used in the method, which are proteases that are well known to one of skill in the art. Thus, this is not a “laundry list disclosure of every moiety,” but rather a disclosure of the actual scaffold proteases contemplated for mutagenesis for use in the method. By virtue of practice of the method, a protease mutein is identified that has increased cleavage activity and/or substrate specificity for a target protein involved in a disease or pathology and inactivates an activity of the disease or pathology.

4) *The Examiner’s statement that Written Description is Not Satisfied Because “only a single protease, granzyme modified at specific positions for mutein identification is described in the working example” does not comport with the Written Description Requirement*

The Examiner is reminded that satisfaction of the written description requirement does not require Working Examples. This basic dogma of written description jurisprudence was recently affirmed in the Federal Circuit’s decision in *Centocor v. Abbot*, whereby the Court stated “we have repeatedly indicated that the written description requirement does

**not demand either examples or an actual reduction to practice”** *Centocor Ortho Biotech, Inc. v. Abbot Labs.*, No. 2010-114, 20-21 (Fed. Cir. 2/23/2011).

As described in detail above, and in previous responses, the specification provides extensive description of all steps of the method as claimed, including description of the step of producing a protease mutein library that contains mutations compared to a scaffold protease. The specification, and claims, describe the particular serine protease scaffolds that are mutagenized, describe the specificity determinant residues that are mutated (*i.e.* between 58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering), and describes methods of mutagenesis. The protease scaffold is merely being used as a scaffold to generate libraries of protease mutants. Indeed, the specificity determinants that are mutated in the library of protease muteins correlate the structure and function of the protease muteins to altered cleavage activity and/or specificity, since changing residues that are known specificity determinants is associated with altering the activity and specificity of the protease. Accordingly, the specification clearly evidences possession of the claimed subject matter since provision of numerous serine protease scaffolds and identification of specificity determinant residues to be mutated, permits one of skill in the art to clearly envision and visualize the library of protease muteins used in the practice of the method.

**V. THE REJECTION OF CLAIMS 1-7, 9, 11-16, 45, 48, 51-53, 56-59, 61-63 AND 65-78 UNDER 35 U.S.C. §102**

Claims 1-7, 9, 11-16, 45, 48, 51-53, 56-59, 61-63 and 65-78 are rejected under 35 U.S.C. §102(b) as being anticipated by Lien *et al.* (Combinatorial Chemistry and High Throughput Screening, 1999)(as evidenced by Shi *et al.*, USP 20020197701). The Examiner states that Lien *et al.* discloses a method of identifying serine proteases using targeted combinatorial mutagenesis of serine proteases with N mutations by generating a library of mutant proteases, contacting the library with a substrate and identifying the mutant. For example, the Examiner states that Lien *et al.* discloses “quantitative assessments of cleavage made by monitoring the hydrolysis (inactivation as claim)...” and further states that Lien *et al.* discloses other various limitations of the claimed subject matter. The Examiner further states that Lien *et al.* discloses the same effect achieved by the specific proteases, *i.e.* inactivation or cleavage of the target protein, and that Lien *et al.* at page 73, col. 1 discloses that it is useful to generate proteases with new or desirable cleavage specificities, such as for



use in practical applications in biotechnology such as in therapy for the modulation of zymogen cascades, such as blood coagulation.

This rejection is respectfully traversed. The arguments at pages 31-39 provided with the response mailed July 20, 2010 are incorporated by reference herein.

#### **Relevant Law**

The case law related to anticipation is set forth in previous responses, incorporated by reference herein.

#### **The Rejected Claims**

The rejected claims are set forth above. It is noted that, as amended, **all claims** recite that in the step of producing a library of protease muteins containing mutations of a serine protease scaffold or a particular recited serine protease scaffold, **a mutation is at an amino acid position selected from among amino acid residue between 58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering.**

#### **Analysis**

The disclosure of Lien *et al.*, and the distinction of the claims from the disclosure of Lien *et al.* has been discussed in the previous response mailed July 20, 2010, incorporated by reference herein. The Examiner is reminded that anticipation requires the disclosure in a single prior art reference of each element of the claim under consideration. In this instance, Lien *et al.* does not disclose a method using a protease library, whereby the library contains protease muteins having mutations in a serine protease scaffold at an amino acid residue **between 58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering.** Hence, Lien *et al.* does not anticipate the claimed subject matter at least for this reason.

Notwithstanding this, in addition, it respectfully is submitted that there is no disclosure in Lien *et al.*, alone or as evidenced by Shi *et al.*, of a method of identifying a protease mutein that cleaves and **inactivates** an activity of a target protein involved in a disease or pathology. The Examiner appears to be equating cleavage and hydrolysis with inactivation of an activity of a protein. **This is not correct.** All proteases hydrolyze peptide bonds upon cleavage of a target protein, yet not all proteases inactivate an activity of the target protein. Hydrolysis of peptide bonds results in cleavage of a target; it does not describe whether the cleaved protein is inactive or active. In fact, cleavage or products can include activation cleavage or inactivation cleavage. In biological processes many proteases **activate**

an activity of a target protein. This occurs, for example, in proteolytic cascades, such as is involved in complement activation, coagulation and apoptosis, whereby the protease acts as a catalyst to **activate** downstream targets. For example, thrombin cleaves by hydrolysis the target protein substrates involved in the coagulation pathway. In particular, the downstream activated targets of thrombin are described in Bode *et al.* (1992) Protein Science, 1:426:471 as follows:

$\alpha$ -Thrombin converts fibrinogen into fibrin, which consequently aggregates and forms the interconnecting network of thrombi. Furthermore, thrombin is able to **activate** several coagulation and plasma factors, such as Factor V, Factor VII, XIII, and protein C. [emphasis added]

With respect to Lien *et al.*, Lien *et al.* discloses the use of targeted combinatorial mutagenesis (TCM) in the screening methods to engineer proteases. In particular, Lien *et al.* reviews the results of various methods in terms of elucidating the roles of various residues in altering the cleavage activity or substrate specificity of the protease. Lien *et al.* summarizes one study (Tsiang *et al.* (Biochemistry (1996) 3:16449) that used random substitution, instead of TCM, to alter the substrate preference of thrombin from fibrinogen to protein C (see *e.g.* last sentence of first paragraph on page 73). As noted above, naturally, **activation** of protein C and fibrinogen is mediated by cleavage by thrombin. Hence, according to Lien *et al.*, the study identified a mutant protease with preferential protein C **activation** over fibrinogen activation. Thus, while Lien *et al.* discloses that such a thrombin mutant can be used in therapy for modulation of zymogen cascades, such as blood coagulation, the resulting mutant summarized in the review article of Lien *et al.* **has the opposite effect** of the protease mutants identified by the instant method. There is no disclosure in Lien *et al.* of any method of identifying a protease that inactivates an activity of a target protein, and in particular a protein involved in a disease or pathology. Hence, Lien *et al.* also fails to disclose this limitation of the claims, and for this reason, also does not anticipate the instant claims.

#### **Rebuttal to Examiner's Comments**

1) The Examiner states that the limitation in claim 1 of the recitation of target proteins, wherein "the target protein is selected from among a cell surface molecule that transmits an extracellular signal for cell proliferation, a cytokine; a cytokine receptor and a signaling protein that regulates apoptosis" is not accorded any weight because it is in the preamble. The Examiner reasons that the preamble is used only to state a purpose or intended use for the target protein. This is not correct. First, the limitation of the target protein is not in the preamble, since it follows the transition clause "wherein" in the claim. Hence, the

limitation is part of the claim. Nevertheless, in the interest of advancing this case to allowance and for clarity, the independent claim 1, 53, 59 and 63 are amended herein to move the limitations in the claim that are recited above "comprises" to below the transition phrase "comprises" under the particular steps of the method. Hence, the recitation of the particular target protein or the protease used in the method are explicit limitations in the claims and should be considered and given weight when reviewing the claim in view of cited references.

Second, the limitation is not made to state a purpose or intended use for the target protein but rather limits the particular target protein or substrate sequence of the target protein that is used in the method. In particular, the limitation of "a cell surface molecule that transmits an extracellular signal for cell proliferation, a cytokine, a cytokine receptor and a signaling protein that regulates apoptosis" are known classes or proteins. Table 1 in the specification at page 9, and as reproduced above, provides exemplary known target proteins that are representative of these classes of proteins.

2) The Examiner states that Lien *et al.* discloses a truncated human growth hormone substrate for proteolytic cleavage, which is included in claim 63, and cites to the paragraph bridging pages 86-87 of Lien *et al.* First, Applicant respectfully submits that growth hormone is not recited as a target protein in claim 63, and hence Lien *et al.* cannot anticipate claim 63 for this reason. Also, notwithstanding the fact that Lien *et al.* does not disclose the other limitations of the claims, Applicant can find no reference in Lien *et al.* of a method of screening a protease mutein for cleavage activity and/or specificity for a target substrate that is a growth hormone. The only reference to growth hormone on page 87 of Lien *et al.* refers to the use of phage display to engineer human growth hormone (citing to Lowman *et al.* (1993) J Mol Biol, 234:564-78). This study is completely different from the instant method because it is directed to the affinity maturation of growth hormone for binding to its receptor. It has nothing to do with a method of screening protease muteins against a target substrate, in particular a target substrate involved in a disease or pathology.

#### **VI. THE REJECTION OF CLAIMS 1-5, 7, 9, 11, 13-16, 48, 51-53, 58-59, 61-63, 65 AND 67 UNDER 35 U.S.C. §103**

Claims 1-5, 7, 9, 11, 13-16, 48, 51-53, 58-59, 61-63, 65 and 67 are rejected under 35 U.S.C. §103(a) as being unpatenable over Guinto *et al.* (Proc. Natl. Acad. Sci. USA, 96, pp.1852-1857, March 1999). The Examiner states that Guinto *et al.* teaches method of identifying a mammalian protease from a database a total of 284 enzymes. In particular, the Examiner states that the variability of amino acid residue at position 225 is assessed, and that

mutation of this residue is identified as influencing the catalytic activity of the enzyme.

Further, the Examiner states:

The claim ability of the candidate mutant as a therapeutic for treatment of the disease associated with the target substrate is a property considered inherent to the prior art mutant serine proteases. The ability of the mutant protease to catalyze the target protein such as fibrinogen or protein C will result in the inactivation of the target hence its amelioration of the disease associated with fibrinogen or protein C.

Applicant respectfully traverses this rejection. It is noted that claim 6, 12, 45, 56 (now cancelled), 57, 66 and 68-79 are not rejected on this ground.

#### **Relevant Law**

The case law related to obviousness is set forth in previous responses filed August 7, 2008, June 11, 2009 and July 20, 2010, each incorporated by reference herein.

In addition, the concept of inherency is not applicable to the question of obviousness. In *re Sporman*, 363 F.2d 444, 150 USPQ 449 (CCPA 1965). To state that a reference inherently teaches an element of a claim rebuts any finding of *prima facie* obviousness. The concept of inherency is not properly applicable to the question of obviousness (see, *In re Sporman*, 363 F.2d 444, 150 USPQ 449 (CCPA 1965)).

Obviousness and inherency are entirely different questions; that which may be inherent is not necessarily known and, therefore, is an indication of unobviousness (*In re Sporman*, 363 F.2d 444, 449, 150 USPQ 449, 452 (CCPA 1965; see, also *In re Naylor*, 360 F.2d 765, 152 USPQ 106 (CCPA 1966); *In re Adams*, 356 F.2d 998, 148 USPQ 742 (CCPA 1966); and *In re Shetty*, 566 F.2d 81, 195 USPQ 753 (CCPA 1977)).

Reference to an unexpected property as inherent begs the question of whether an unexpected property rebuts *prima facie* obviousness. The unexpected property is part of the invention as a whole, and, therefore, evidence of unobviousness of the claimed subject matter. In *In re Naylor* a process for preparing a polybutadiene polymer having unexpected properties was at issue. The CCPA held that the fact that a rubbery polybutadiene having high 1,2-addition might be inherent in following the combined teachings of the prior art is immaterial, if one of ordinary skill in the art would not appreciate or recognize that inherent result. In *In re Adams*, the CCPA held that since properties of a claimed structure are always inherent, it is "transparently erroneous" to state that subject matter cannot be patented on the basis of an inherent property. In *re Shetty*, the court held that "inherency is quite immaterial if, as record established here, one of ordinary skill in the art would not appreciate or recognize that inherent result."



### The Rejected Claims

The rejected claims are set forth above. It is noted that, as amended, **all claims** recite that in the step of producing a library comprising protease mutein, the protease scaffold, which in all claims is a serine protease or a particular recited serine protease, **comprise a mutation at an amino acid position selected from among amino acid residue between 58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering.**

### Analysis

Guinto *et al.* teaches the role of the amino acid residue 225, numbered based on chymotrypsin numbering like the instant application, in serine proteases. Guinto *et al.* teaches studies involving saturation mutagenesis of thrombin at position 225. The resulting mutants were then tested for their ability to hydrolyze a chromogenic substrate, fibrinogen and protein C, as well as their ability to be inhibited by the inhibitor antithrombin II. Guinto *et al.* teaches that the mutants have reduced catalytic activity.

First, as noted above, the Examiner's reasoning that inactivation of the target is inherent to the mutant is flawed. First, the Examiner, who states that the ability of the candidate mutant as a therapeutic for treatment of the disease associated with the target substrate is a property considered inherent to the prior art mutant serine proteases, is reminded (see discussion of case law above) that inherency has no place in a consideration of obviousness. Inherent properties are relevant in the context of anticipation, but in the context of an obviousness rejection, such properties can evidence unexpected results. As discussed below and in responses of record, no prior art of record before the earliest priority date of the instant application teaches or suggests any method of engineering or identifying a protease mutein that inactivates an activity of a target protein involved in a disease or pathology, nor that protease muteins identified by such a method could be candidate therapeutics.

More importantly, the Examiner's reasoning is flawed, since hydrolysis of a target protein is not synonymous with inactivation of an activity of the target protein. As evidenced by Bode *et al.* ((1992) Protein Science, 1:426:471) discussed above, thrombin hydrolyzes and cleaves its target substrates to **activate** them. **Hence, inactivation of the target protein substrates by thrombin is not an inherent activity of the protease.** Rather, the inherent activity of thrombin is that it activates its protein target substrates. Applicant notes that the instant claims include an affirmative step of testing the protease mutein for inactivation of an activity of the protease. It respectfully is submitted that a method of engineering or screening

for a protease mutein, including a protease mutein of thrombin, that inactivates an activity of a target protein substrate is not taught or suggested in Guinto *et al.* There is no evidence in Guinto *et al.* of any protease that hydrolyzes and inactivates an activity of a protein target substrate, nor any teaching or suggestion of a method of screening to identify new protease muteins engineered to have this activity towards particular target proteins involved in a disease or pathology.

Notwithstanding this, Guinto *et al.* only is directed to elucidation and mutagenesis of the amino acid residue 225. There is no teaching or suggestion of any other residues that are specificity determinants in serine proteases that, when mutated, can alter the cleavage activity and/or substrate specificity towards a target protein involved in a disease or pathology. In particular, there is no teaching or suggestion that any of amino acid residues **58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering**, are specificity determinants. Hence, there is no teaching or suggestion in Guinto *et al.* that a method of screening a library of protease muteins containing mutations at any of the above residues could be used to identify a protease mutein that has an increased cleavage activity and/or substrate specificity for a target protein involved in a disease or pathology, and that by virtue of that is identified as inactivating an activity of the target protein.

In sum, Guinto *et al.* fails to render the claims obvious because it does not teach or suggest all limitations of the claims. For example, Guinto *et al.* does not teach or suggest producing a library of protease muteins that have mutations at amino acid residues 58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering. Guinto *et al.* also does not teach or suggest a step of testing an identified protease mutein for inactivation of an activity of the target protein. Since a *prima facie* case of obviousness requires that the reference teach or suggest all elements as claimed, the Examiner has failed to set forth a case of *prima facie* obviousness.

#### **Rebuttal to Examiner's Comments**

1) In reply to Applicant's argument in the last response that the target proteins cleaved by the mutant thrombins are activated and not inactivated, the Examiner state in reply:

Guinto teaches above inhibitors, *i.e.*, inactivation of the target protein as claimed. The resultant effect of said inhibition would inherently ameliorate a disease or pathology to which the target protein is involved. The function of the target protein involvement with the numerous diseases is irrelevant to the screening protease that binds or inhibits said target protein.

It is respectfully submitted that it is not clear to Applicant what “inhibitors” the Examiner is referring to in the teachings of Guinto *et al.* Nevertheless, in the arguments above and in the last response mailed July 20, 2010, Applicant stated that Guinto *et al.* does not teach that the protease mutants mutated at position 225 inactivate an activity of any target protein. Rather, thrombin, and the thrombin mutants as taught in Guinto, activate the target substrates; Guinto *et al.* teaches that mutation of residue 225 results in reduced catalytic activity and presumably reduced ability to activate the target substrates. Since inactivation of a protein target substrate is not an inherent activity of hydrolysis by a protease, the Examiner has no basis to conclude that Guinto *et al.* inherently teaches inhibition of target proteins. There is no teaching or suggestion in Guinto *et al.* that hydrolysis and cleavage of a thrombin target protein, such as protein C, results in inactivation of an activity of protein C. Guinto *et al.* does not teach the thrombin protease mutants are inhibitors, nor does it teach that the thrombin protease mutants inactivate an activity of any of its target substrates.

2) In reply to Applicant’s argument in the last response that Guinto *et al.* fails to disclose any of the recited target proteins in the claims that are a “cell surface molecule that transmits an extracellular signal for cell proliferation, a cytokine, a cytokine receptor and a signaling protein that regulates apoptosis,” the Examiner again refers the Applicant to the previous comment about the preamble. In reply to this, Applicant refers the Examiner to the Rebuttal 1) above on page 28.

**VII. THE REJECTION OF CLAIMS 1-7, 9, 11-16, 45, 48, 51-53, 56-59, 61-63, AND 65-78 UNDER 35 U.S.C. §103**

Claims 1-7, 9, 11-16, 45, 48, 51-53, 56-59, 61-63, and 65-78 are rejected under 35 U.S.C. §103(a) as being unpatenable over Lien *et al.* (Combinatorial Chemistry and High Throughput Screening, 1999, 2:73-90) in view of either Harris *et al.* I (J. Biol. Chem., 273:27364-27373 (1998) or Harris *et al.* II (Current Opinion in Chemical Biology, 2:127-132 (1998)) and Waugh *et al.* (Nature Structure Biology). The Examiner states that while Lien *et al.* does not disclose the enzyme as granzyme and the substrate as caspase, Harris *et al.* I teaches the substrate specificity of granzyme B and that caspases are likely substrates, and teaches that Arginine 192 is a structural determinant of specificity of granzyme B; Harris *et al.* II teaches the same method as Harris *et al.* I; and Waugh *et al.* teaches that granzymes are a vital component of the ability to induce apoptosis. Thus, the Examiner concludes that it would have been obvious to one of ordinary skill to use the serine protease granzyme and the target protein caspase in the method of Lien *et al.*

Applicant respectfully traverses this rejection. This rejection is similar to previous rejections made of record that have been addressed in responses filed August 7, 2008, June 11, 2009 and July 20, 2010, each of which are incorporated by reference herein.

#### **Relevant Law**

The case law related to obviousness is set forth in previous responses filed August 7, 2008, June 11, 2009 and July 20, 2010, each incorporated by reference herein.

#### **The Rejected Claims**

The rejected claims are set forth above. It is noted that, as amended, **all claims** recite that the step of producing a library of protease mutein, the protease muteins in the library contain mutations in a serine protease scaffold or in a particular recited serine protease scaffold **at an amino acid position selected from among amino acid residue between 58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering.**

#### **Analysis**

As discussed above and in previous responses of record, Lien *et al.* does not teach or suggest a method including all limitations as instantly claimed. For example, Lien *et al.* does not teach or suggest a method using a protease library, whereby the library contains protease muteins having mutations in a serine protease scaffold at an amino acid residue **between 58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering.** Lien *et al.* also does not teach or suggest any method of identifying a protease that inactivates an activity of a target protein involved in a disease or pathology, since contrary to the Examiner's comments throughout this response, inactivation of an activity of a target protein is not inherent to hydrolysis of that protein by a protease. None of Harris *et al.* I, Harris *et al.* II or Waugh *et al.*, singly or in any combination, cures these defects.

Harris *et al.* I is directed to a study of the mechanism by which granzyme B mediates apoptosis of target cells and demonstrates that granzyme B displays extended substrate specificity and proposes a model by which granzyme B acts. In this context, for example, while Harris *et al.* I teaches two mutants of granzyme B at amino acid residue 192, Harris *et al.* teaches that mutation of amino acid residue 192 in granzyme B **reduces** its hydrolysis of an optimal tetrapeptide substrate. Harris *et al.* does not teach or suggest employing granzyme B nor any protease as a scaffold in a method for generating and identifying



proteases that inactivate an activity of a target protein to produce a candidate therapeutic. Thus, Harris *et al.* is of no relevance to the instant claims.

Further, there is no teaching or suggestion in Harris *et al.* I of any protease mutants that exhibit increased cleavage activity and/or substrate specificity for a target protein involved in a disease or pathology as required by the instant claims. Indeed, not only does Harris *et al.* teach that the granzyme B mutants exhibit **reduced** hydrolysis of the synthetic substrate, the synthetic substrate is **not** a target protein involved in a disease or pathology, in particular any target protein recited in the instant claims. Also, Harris *et al.* I does not teach or suggest any method that includes as a step testing the identified protease mutant for inactivation of an activity of the target protein. In this regard, Harris *et al.* teaches that caspase is a target substrate of granzyme B. Like the coagulation pathway, the apoptosis pathway involves a proteolytic cascade, whereby proteases act as catalysts to activate downstream molecules. For example, granzymes such as granzyme B cleave caspases, thereby altering the conformation and leading to their *activation*. Specifically, granzyme B cleaves procaspase-3 to its activated caspase-3 form, which itself performs further downstream functions in the apoptosis pathway. For example, Harris *et al.* I teaches at page 27364, 2<sup>nd</sup> column:

Although granzyme B is the only known mammalian serine protease to have P1-proteolytic specificity, it is shared with the caspases, a family of cysteine proteases that are also activated during apoptosis. The link between granzyme B and the caspases has been strengthened by studies indicating that **granzyme B can cleave and activate certain members of the caspases**, and it has been suggested that this is one of the mechanisms by which granzyme B mediates apoptosis in vivo. [emphasis added]

Harris *et al.* II is even of less relevance to the instant claims. Harris *et al.* II broadly teaches methods of engineering the substrate specificity of proteases. Harris *et al.* does not identify any of residues 58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering as specificity determinants in serine proteases, nor suggest their modification. Hence, Harris *et al.* II does not teach or suggest a method using a protease library, whereby the library contains protease muteins having mutations in a serine protease scaffold at an amino acid residue **between 58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering**. Harris *et al.* II also does not teach or suggest any method of identifying a protease that inactivates an activity of a target protein involved in a disease or pathology. Thus, Harris *et al.* II, singly or

in combination with Lien *et al.* or Harris *et al.* I, does not cure the deficiency in rendering the claims *prima facie* obvious.

Finally, Waugh *et al.* also is of little relevance to the instant claims. Waugh *et al.* teaches that granzymes are involved in inducing apoptosis by acting on downstream substrates such as caspases by activation cleavage. Waugh *et al.* also teaches the elucidation of the molecular determinants of substrate specificity in granzyme B. In particular, the first paragraph of Waugh *et al.* teaches:

Granzymes are a vital component of the cytotoxic lymphocyte's ability to induce apoptosis, contributing to rapid cell death of a tumor or virally infected target cell by the cleavage of downstream substrates and the **activating** cleavage of caspases. [emphasis added]

Thus, Waugh *et al.* is directed to the normal function of granzyme to activate caspases to induce apoptosis. Waugh *et al.* does not teach or suggest any protease that inactivates an activity of a target protein, in particular a target protein involved in a disease or pathology, nor a method of identifying such a protease mutant as instantly claimed. Further, while Waugh *et al.* teaches the elucidation of the molecular determinants of substrate specificity in granzyme B, Waugh *et al.* does not teach or suggest that mutation of any of these residues would alter substrate specificity, in particular substrate specificity for a target protein involved in a disease or pathology. Waugh *et al.* does not teach or suggest any mutants of granzyme B. There is no teaching or suggestion in Waugh *et al.* of a method of generating mutants of granzyme B, such as a library of protease mutants as claimed that contain mutations at an amino acid residue between 58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering. There is no teaching or suggestion in Waugh *et al.* of any method of screening for mutant proteases to identify a protease that has increased cleavage activity and/or substrate specificity for a target protein involved in a disease or pathology that includes a step of testing the identified protease mutants for inactivation of an activity of the target protein. Since Waugh *et al.* also does not teach or suggest a library of protease mutants at any of the noted residues that are specificity determinants, and does not teach or suggest testing any mutants for inactivation of an activity of a target protein involved in a disease or pathology, Waugh *et al.*, singly or in combination with Lien *et al.*, Harris *et al.* I or Harris *et al.* II, does not teach or suggest the claimed method.

Hence, the Examiner has failed to set forth a *prima facie* case of obviousness because none of Lien *et al.*, Harris *et al.* I, Harris *et al.* II and/or Waugh *et al.*, singly or in any

combination thereof, teaches all elements as claimed. Each of independent claims 1, 53, 59 and 63 is directed to a method of producing and identifying a protease mutein that cleaves and inactivates a target protein involved in a disease or pathology, including steps of: 1) producing a library of protease muteins each having N mutations at an amino acid residue between 58 and 64, 97, 98, 99, 100, 171, 174, 180, 189, 190, 191, 192, 215, 217, 218 and 226 by chymotrypsin numbering, relative to a particular recited serine protease scaffold; 2) measuring the cleavage activity and/or substrate specificity of at least two members in the library for a substrate sequence in a target protein involved in a disease or pathology; 3) identifying those protease muteins in the library that have increased cleavage activity and/or altered substrate specificity relative to a wild-type scaffold protease; and 4) testing the identified proteases for cleavage and inactivation of an activity of the target protein that contains the substrate sequence to identify a protease mutein or a biologically active portion thereof that inactivates an activity of a target protein that is involved with the disease or pathology. In particular, the cited references, singly or in combination, do not teach or suggest a method that includes a step of producing a library of protease muteins containing mutations at the recited specificity determinants, measuring and identifying protease muteins that have increased cleavage activity and/or substrate specificity for a substrate sequence in a target protein or a target protein involved in a disease or pathology, or a step of testing identified protease muteins for inactivation of an activity of the target protein. Since *prima facie* obviousness requires that the references, singly or in any combination, teach or suggest all elements as claimed, the Examiner has failed to set forth a case of *prima facie* obviousness.

#### **Rebuttal to Examiner's Comments**

1) In reply to Applicant's argument in the previous response that Harris *et al.* teaches that granzyme B **activates** caspase, and thus does not teach or suggest a method including a step of testing for inactivation of an activity of the target protein, the Examiner points towards the portion of claim 1 that recites "regulates" apoptosis and states that this encompasses "activation." The Examiner continues, however, to state that the limitation in claim 1 of the target protein is in the preamble and is not given any weight.

First, Applicant's argument that Harris *et al.* teaches that granzyme B activates caspase goes to the limitation in the claims of step e) "testing the identified protease(s) or biologically active portion thereof for cleavage and inactivation of an activity of the target protein that contains the substrate sequence, thereby identifying a protease mutein or a biologically active portion thereof that inactivates an activity of a target protein that is involved with the disease

or pathology.” Harris *et al.* I, which only is directed to a study of the function of granzyme B to activate caspase, does not teach or suggest any method for screening proteases, including any method that includes a step of testing or identifying a protease that inactivates an activity of a target protein. There is no teaching or suggestion in Harris *et al.* I for preparing a library that contains mutant granzyme B enzymes, nor that granzyme B could ever be engineered to inactivate an activity of any target protein.

For this reason, the Examiner’s reference to the portion of claim 1 that recites “regulates apoptosis” is misplaced. This portion of the claim is a limitation on the particular target protein that is used in the method. For example, a caspase is a target protein that regulates apoptosis and could be used in the instant method. The limitation on what the target protein is, is not relevant to the endpoint of the method of testing and identifying a protease mutein for inactivation and activity of that target protein.

Further, as noted above, the claims are amended herein to render it clear that the limitation on the target protein is not part of the preamble, but are explicit limitations of the claimed method. Applicant respectfully requests that the Examiner consider and give all recited limitations weight.

2) The Examiner refers to the Examples to state that the findings of the instant application are the same as Harris *et al.*, since the instant application also shows that an exemplified protease mutein activates caspase-3. For example, the Examiner cites page 53, lines 18-20 of the instant application, which states: “at low concentrations the mutant activates caspase-3 by cleaving at the activation sequence (SEQ ID NO:4).

***It respectfully is submitted that the Examiner must review the results in the Example as a whole.*** Picking and choosing only portions of the Example is not representative of the results depicted in the Example. In this case, a complete review of the results in Example 11 exemplifies that the exemplified protease granzyme B mutein I99A/N218A **inactivates** the activity of caspase by cleavage at an identified inactivation sequence at residues 260-265 (SEQ ID NO:2) (see *e.g.* at page 52, lines 6-7). In contrast to cleavage of the activation sequence that is set forth in SEQ ID NO:4, which is the normal sequence cleaved by wild-type granzyme B to activate caspase, cleavage of this inactivation sequence in caspase-3 set forth in SEQ ID NO:2 inactivates an activity of caspase-3. Thus, the Example compares the activity of wildtype granzyme B, (which normally cleaves that activation sequence in caspase-3 set forth in SEQ ID NO:4) and the mutein I99A/N218A for inactivation cleavage and inactivation of an activity of caspase-3.



The Example states that the mutant granzyme B I99A/N218A is able to cleave caspase-3 at the inactivation sequence as represented by a shift in cleavage product of the appropriate size for the cleaved peptide, while “wild-type granzyme B does not cleave the peptide” (see *e.g.* at page 52, lines 16-27). The Example further describes that the cleavage of the inactivation sequence results in cleavage of the small subunit of caspase-3 (see *e.g.* at page 52, line 28 to page 53, line 2). Further, the Example describes that the mutant can cleave and inactivate an activity of full-length caspase-3 as assessed by assaying the activity of caspase-3 for its substrate Ac-DEVD-AMC (see *e.g.* at page 53, lines 3-7). The results in the Example show that the mutant granzyme B dramatically inactivates the activity of caspase-3 with a Vmax in the presence of the mutant granzyme B of approximately zero (see *e.g.* at page 53, lines 7-13). The Example further describes that the mutant granzyme B also inhibits an activity of caspase-3 as assessed in assays assessing apoptosis. While the Example shows that at low concentrations the mutant activates caspase-3 by cleaving the activation sequence, the result states that “at high concentrations it inhibits caspase-3 by cleaving at the inactivation sequence” (see *e.g.* at page 53, lines 14-24). Finally, the Example demonstrates that the mutant granzyme B can effectively antagonize caspase-induced activation by wildtype granzyme B (see *e.g.* at page 53, line 25 to page 54, line 3). For example, the Example states that “as shown in Figure 7B, the mutant granzyme B antagonized the effect of wildtype granzyme B to induce apoptosis by inactivating the caspase-3.”

Thus, the Examples clearly demonstrates practice of the method as claimed by generating a protease mutant; measuring the cleavage activity and/or substrate specificity for a substrate sequence in a target protein (here, the inactivation cleavage sequence in caspase-3 as set forth in SEQ ID NO:2) and identifying a mutant that has increased cleavage or substrate specificity therefor; and testing the identified mutant for inactivation of an activity of the target protein (here, the activity of caspase-3 to cleave is substrate Ac-DEVD-AMC or the activity of caspase-3 to induce apoptosis). This is in direct contrast to the findings in Harris *et al.*

#### **VIII. THE REJECTION OF CLAIMS 1-7, 9, 11-16, 45, 48, 51-53, 56-59, 61-63, AND 65-78 FOR NONSTATUTORY DOUBLE PATENTING**

Claims 1-7, 9, 11-16, 45, 48, 51-53, 56-59, 61-63, and 65-78 are provisionally rejected on the ground of nonstatutory obviousness-type double patenting as being unpatentable over claims 1-10 of copending Application No. 12/005,949 ('949) application.

Applicant : Nguyen *et al.*  
Serial No. : 10/677,977  
Filed : October 02, 2003

Attorney Docket No.: 33328.04905.US01/4905  
Amendment and Response

As stated in the response mailed July 20, 2010 and previous responses, Applicant requests deferral of resolution of this issue. It is not possible to assess whether claims at allowance in each application will overlap requiring a terminal disclaimer until there is an indication of allowable subject matter in at least one application. It is premature to file a terminal disclaimer at this time. If, when one or both applications is deemed allowed, it is determined that a terminal disclaimer is necessary, Applicant will file a terminal disclaimer.

\* \* \*

Consideration of the above remarks, entry of this amendment and continued examination of the application on the merits respectfully are requested.

Respectfully submitted,

---

Stephanie Seidman  
Reg. No. 33,779

Attorney Docket No. 33328.04905.US01/4905  
**Address all correspondence to: 13565**  
Stephanie Seidman  
Mckenna Long & Aldridge LLP  
4435 Eastgate Mall, Suite 400  
San Diego, California 92122  
Telephone: (619) 595-8010  
Facsimile: (858) 595-8135  
email: sseidman@mckennalong.com

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

Applicant	: Nguyen <i>et al.</i>	Art Unit	: 1639
Serial No.	: 10/677,977	Examiner	: Wessendorf, Teresa D.
Filed	: October 02, 2003	Conf. No.	: 9061
Cust. No.	: 77202		
Title	: <b>METHODS OF GENERATING AND SCREENING FOR PROTEASES WITH ALTERED SPECIFICITY</b>		

**APPENDIX**

1. Izban et al., "Characterization of the interleukin-1 beta-converting enzyme/ced-3-family protease, caspase-3/CPP32, in Hodgkin's disease: lack of caspase-3 expression in nodular lymphocyte predominance Hodgkin's disease," *Am J Path* 154(5):1439-1447 (1999)
2. Rubin et al., "Purification and characterization of a human tumor necrosis factor from the LuKII cell line," *Proc Natl Acad Sci USA* 82:6637-6641 (1985)
3. Williamson et al., *Proc Natl Acad Sci USA* 80:5397-5401 (1983)
4. Lachman et al., *J. Immunol.* 119:2019 (1977)
5. Howard et al., "Role of interleukin 1 in anti-immunoglobulin-induced B cell proliferation," *J. Exp Med* 157:1529-1543 (1983)
6. Lowenthal and MacDonald, "Binding and internalization of interleukin 1 by T cells," *J. Exp Med* 164:1060-1074 (1986)
- 7.
8. Gillis and Mizel, "T-Cell lymphoma model for the analysis of interleukin 1-mediated T-cell activation," *Proc Natl Acad Sci USA* 78(2):1133-1137 (1981)
9. Grabstein et al., "Purification to homogeneity of B cell stimulating factor. A molecule that stimulates proliferation of multiple lymphokine-dependent cell lines," *J Exp Med*, 163:1405-1414 (1986)
10. Kammer et al., "Homodimerization of interleukin-4 receptor  $\alpha$  chain can induce intracellular signaling," *J Biol Chem* 271(39):23634-23637 (1996)
11. Aoki et al., "Interleukin 5 enhances interleukin 2-mediated lymphokine-activated killer activity," *J Exp Med*, 170:583-588 (1989)
12. Allavena et al., "Interleukin-12 is chemotactic for natural killer cells and stimulates their interaction with vascular endothelium," *Blood* 84:2261-2268 (1994)
13. Caput et al., "Cloning and Characterization of a Specific Interleukin (IL)-13 Binding Protein Structurally Related to the IL-5 Receptor  $\alpha$  Chain," *J Biol Chem* 271(28):16921-16926 (1996)

14. Celi et al., "P-selectin induces the expression of tissue factor on monocytes," *Proc Natl Acad Sci USA* 91:8767-8771 (1994)
15. Moore et al., "The P-selectin Glycoprotein Ligand from Human Neutrophils Displays Sialylated, Fucosylated, O -Linked Poly-N-acetyllactosamine," *J Biol Chem* 269:23318-23327 (1994)
16. Payan et al., "Substance P recognition by a subset of human T lymphocytes," *J Clin Invest.* 74:1532-1539 (1984)
17. Thayer et al., "Regulation of calcium homeostasis in sensory neurons by bradykinin," *J Neuroscience* 8(11):4089-4097 (1988)
18. Chung et al., "purification and characterization of an abnormal Factor IX (Christmas factor) molecule," *J Clin Invest* 62:1078-1085 (1978)
19. Picard et al., "Multiple extracellular domains of CCR-5 contribute to human immunodeficiency virus type 1 entry and fusion," *J Virology* 71:5003-5011 (1997)
20. Brelot et al., "Role of the first and third extracellular domains of CXCR-4 in human immunodeficiency virus coreceptor activity," *J Virology* 71:4744-4751 (1997)
21. Willey et al., "In vitro mutagenesis identifies a region within the envelope gene of the human immunodeficiency virus that is critical for infectivity," *J Virology* 62:139-147 (1988)
22. Gabuzda et al., "Effects of deletions in the cytoplasmic domain on biological functions of human immunodeficiency virus type 1 envelope glycoproteins," *J Virology* 66:3306-3315 (1992)
23. Sekikawa and Lai, "Defects in functional expression of an influenza virus hemagglutinin lacking the signal peptide sequences," *Proc Natl Acad Sci USA* 80:3563-3567 (1983)
24. Azuma et al., "CD28 interaction with B7 costimulates primary allogeneic proliferative responses and cytotoxicity mediated by small, resting T lymphocytes," *J Exp Med* 175:353-360 (1992)
25. Pages et al., "Two distinct intracytoplasmic regions of the T-cell adhesion molecule CD28 participate in phosphatidylinositol 3-kinase association," *J Biol Chem* 271:9403-9409 (1996)
26. Chang et al., "Dissection of the human CD2 intracellular domain. Identification of a segment required for signal transduction and interleukin 2 production," *J Exp Med* 169:2073-2083 (1989)
27. Thali et al., "Effects of changes in gp120-CD4 binding affinity on human immunodeficiency virus type 1 envelope glycoprotein function and soluble CD4 sensitivity," *J Virology* 65:5007-5012 (1991)



28. Suzuma et al., "Vascular endothelial growth factor induces expression of connective tissue growth factor via KDR, Flt1, and phosphatidylinositol 3-kinase-akt-dependent pathways in retinal vascular cells," J Biol Chem 275:40725-70731 (2000)
29. Feige and Baird "Basic fibroblast growth factor is a substrate for protein phosphorylation and is phosphorylated by capillary endothelial cells in culture," Proc Natl Acad Sci USA 86:3174-3178 (1989)
30. Moriai et al., "A variant epidermal growth factor receptor exhibits altered type alpha transforming growth factor binding and transmembrane signaling," Proc Natl Acad Sci USA 91:10217-10221 (1994)
31. Ranges et al., "Inhibition of cytotoxic T cell development by transforming growth factor beta and reversal by recombinant tumor necrosis factor alpha," J Exp Med 166:991-998 (1987)
32. Accession No. P00761
33. Accession No. P07477 Trypsin-1
34. Accession No. P07478 Trypsin-2
35. Accession No. P35030 Trypsin-3
36. Accession No. AAA61231
37. Accession No. AAG30943
38. Accession No. AAC50728
39. Accession No. AAA61232
40. Accession No. P00775
41. Accession No. P35049
42. Accession No. Q99895 Chymotrypsin-C
43. Accession No. P17538 Chymotrypsinogen B
44. Accession No. AAA52128 Chymotrypsinogen B
45. Accession No. AAP36020 Chymotrypsinogen B
46. Accession No. P07338 Chymotrypsin B
47. Accession No. P80646 Chymotrypsin B
48. Accession No. P00768 Chymotrypsin-2
49. Accession No. P00769 Chymotrypsin-2
50. Accession No. O97398
51. Accession No. P55091 Chymotrypsin C
52. Accession No. P04189 Subtilisin E
53. Accession No. AAA22742

28. Suzuma et al., "Vascular endothelial growth factor induces expression of connective tissue growth factor via KDR, Flt1, and phosphatidylinositol 3-kinase-akt-dependent pathways in retinal vascular cells," J Biol Chem 275:40725-70731 (2000)
29. Feige and Baird "Basic fibroblast growth factor is a substrate for protein phosphorylation and is phosphorylated by capillary endothelial cells in culture," Proc Natl Acad Sci USA 86:3174-3178 (1989)
30. Moriai et al., "A variant epidermal growth factor receptor exhibits altered type alpha transforming growth factor binding and transmembrane signaling," Proc Natl Acad Sci USA 91:10217-10221 (1994)
31. Ranges et al., "Inhibition of cytotoxic T cell development by transforming growth factor beta and reversal by recombinant tumor necrosis factor alpha," J Exp Med 166:991-998 (1987)
32. Accession No. P00761
33. Accession No. P07477 Trypsin-1
34. Accession No. P07478 Trypsin-2
35. Accession No. P35030 Trypsin-3
36. Accession No. AAA61231
37. Accession No. AAG30943
38. Accession No. AAC50728
39. Accession No. AAA61232
40. Accession No. P00775
41. Accession No. P35049
42. Accession No. Q99895 Chymotrypsin-C
43. Accession No. P17538 Chymotrypsinogen B
44. Accession No. AAA52128 Chymotrypsinogen B
45. Accession No. AAP36020 Chymotrypsinogen B
46. Accession No. P07338 Chymotrypsin B
47. Accession No. P80646 Chymotrypsin B
48. Accession No. P00768 Chymotrypsin-2
49. Accession No. P00769 Chymotrypsin-2
50. Accession No. O97398
51. Accession No. P55091 Chymotrypsin C
52. Accession No. P04189 Subtilisin E
53. Accession No. AAA22742

Applicant : Nguyen *et al.*  
Serial No. : 10/677,977  
Filed : October 02, 2003

Attorney Docket No.: 33328.04905.US01/4905  
APPENDIX

54. Accession No. AAA22814
55. Accession No. AAA22744
56. Accession No. P00781 Subtilisin DY
57. Accession No. P29599 Subtilisin BL
58. Accession No. P12544
59. Accession No. AAA52647
60. Accession No. AAD00009
61. Accession No. P11032
62. Accession No. P10144
63. Accession No. AAA52118
64. Accession No. AAB59528
65. Accession No. AAA36627
66. Accession No. AAA67124
67. Accession No. P04187
68. Accession No. P51124
69. Accession No. AAA59582
70. Accession No. AAA57262
71. Accession No. Q03238
72. Accession No. O08643
73. Accession No. P32197
74. Accession No. P23946
75. Accession No. AAB26828
76. Accession No. AAA52020
77. Accession No. AAA52021
78. Accession No. AAA52019
79. Accession No. P21842
80. Accession No. P21844
81. Accession No. P56435
82. Accession No. P52195
83. Accession No. P50339
84. Accession No. P00784
85. Accession No. P08246
86. Accession No. AAA36359

Applicant : Nguyen *et al.*  
Serial No. : 10/677,977  
Filed : October 02, 2003

Attorney Docket No.: 33328.04905.US01/4905  
APPENDIX

87. Accession No. AAA36173
88. Accession No. P37357 Neutrophil elastase 2A
89. Accession No. P37358 Neutrophil elastase 2B
90. Accession No. Q9Z284
91. Accession No. P48740 MASP1
92. Accession No. P98064 MASP1
93. Accession No. AAK84071
94. Accession No. Q96P40
95. Accession No. BAB69487
96. Accession No. P08473
97. Accession No. AAA51915
98. Accession No. CAA68752
99. Accession No. AAA52294
100. Accession No. CAA30157
101. Accession No. Q61391
102. Accession No. P08049
103. Accession No. P07861
104. Accession No. P19621
105. Accession No. P09958
106. Accession No. P23188
107. Accession No. Q28193
108. Accession No. P23377
109. Accession No. P29119
110. Accession No. P25779
111. Accession No. AAA30269
112. Accession No. AAA30270
113. Accession No. CAA38278
114. Accession No. AAA30180
115. Accession No. P00749
116. Accession No. AAA61253
117. Accession No. CAA26535
118. Accession No. AAC97138
119. Accession No. P06869



Applicant : Nguyen *et al.*  
Serial No. : 10/677,977  
Filed : October 02, 2003

Attorney Docket No.: 33328.04905.US01/4905  
APPENDIX

120. Accession No. Q8MHY7
121. Accession No. P29598
122. Accession No. Q05589
123. Accession No. P16227
124. Accession No. P04185
125. Accession No. P15120
126. Jonathan Greer (PROTEINS: Structure, Function and Genetics, 7:317-334 (1990))
127. Friedrich *et al.* J Biol Chem. 2002 Jan 18;277(3):2160-8
128. Bode *et al.* (1992) Protein Science, 1:426:471
129. Tsiang *et al.* (Biochemistry (1996) 3:16449
130. Lowman *et al.* (1993) J Mol Biol, 234:564-78

# Characterization of the Interleukin-1 $\beta$ -Converting Enzyme/Ced-3-Family Protease, Caspase-3/CPP32, in Hodgkin's Disease

## *Lack of Caspase-3 Expression in Nodular Lymphocyte Predominance Hodgkin's Disease*

Keith F. Izban,\*<sup>†</sup> Tamara Wrone-Smith,\*<sup>†</sup>  
Eric D. Hsi,<sup>‡</sup> Bertram Schnitzer,<sup>§</sup>  
Maria Eugenia Quevedo,<sup>†</sup> and Serhan Alkan\*<sup>†</sup>

From the Department of Pathology\* and Cardinal Bernardin Cancer Center,<sup>†</sup> Loyola University Medical Center, Maywood, Illinois, the Department of Clinical Pathology,<sup>‡</sup> Cleveland Clinic Foundation, Cleveland, Ohio, and the Department of Pathology,<sup>§</sup> University of Michigan Medical School, Ann Arbor, Michigan

Apoptosis (programmed cell death) serves an important role in the normal morphogenesis, immunoregulation, and homeostatic mechanisms in both normal and neoplastic cells. Caspase-3/CPP32, a member of the ICE/Ced-3-family of cysteine proteases, is an important downstream mediator of several complex proteolytic cascades that result in apoptosis in both hematopoietic and nonhematopoietic cells. Previous studies have demonstrated that caspase-3 is commonly expressed in classical Hodgkin's disease (CHD); however, the biological significance of its expression in Hodgkin's disease is unknown. In this report, the expression of caspase-3 in nodular lymphocyte predominance Hodgkin's disease (NLPHD) was evaluated by immunohistochemistry; in addition, we investigated the role of caspase-3 in CD95 (Fas)-mediated apoptosis in three CHD cell lines. Formalin-fixed, paraffin-embedded tissue sections from 11 cases of NLPHD were immunostained for caspase-3 using a polyclonal rabbit antibody that detects both the 32-kd zymogen and the 20-kd active subunit of the caspase-3 protease. Only 1/11 cases of NLPHD demonstrated caspase-3 immunopositivity in lymphocytic/histiocytic cells. Caspase-3 expression was also evaluated in three CHD cell lines, HS445, L428, and KMH2. Whereas caspase-3 expression was detected in HS445 and L428 cell lines, no expression was found in KMH2 cells by immunohistochemical staining. Treatment of HS445 and L428 cell lines for 72 hours with agonistic CD95 monoclonal antibody induced marked apoptosis that was significantly inhibited by

pretreatment with the caspase-3 inhibitor, DEVD-FMK, as determined by terminal deoxynucleotidyl transferase-mediated dUTP nick end-labeling assay and flow cytometric analysis of 7-amino-actinomycin D staining. In addition, a significant increase in caspase-3 activity as determined by an enzyme colorimetric assay was detected in HS445 and L428 cells after 48 hours of CD95 stimulation. In marked contrast, treatment of caspase-3-deficient KMH2 cells with anti-CD95 mAb did not demonstrate an increase in caspase-3 activity or induce apoptosis. These data demonstrate caspase-3 is important for CD95-mediated apoptosis in CHD cell lines. In addition, the majority of NLPHD cases examined in this study failed to express detectable levels of caspase-3, suggesting these tumor cells may be resistant to apoptotic stimuli dependent on caspase-3 activity. Furthermore, these data suggest the differential expression of caspase-3 noted between NLPHD and CHD may provide additional evidence that each is a unique disease entity. (*Am J Pathol* 1999, 154:1439-1447)

Increased understanding of the physiological and pathological processes of programmed cell death, or apoptosis, at the molecular level will provide insights into carcinogenesis and potentially create new opportunities for development of novel prognostic markers and therapeutic tools for the treatment of various neoplasms. One of the earliest cell death-regulating genes to be identified was the proto-oncogene Bcl-2, an apoptosis inhibitor that appears to block a step in an evolutionarily conserved pathway involved in apoptosis.<sup>1-2</sup> Subsequent investigations led to the isolation of a homologue of Bcl-2 in the nematode *Caenorhabditis elegans*. This homologue, called Ced-9, is necessary for the survival of all cells in this organism.<sup>3</sup> Ced-9 opposes the actions of two cell death-

Accepted for publication February 12, 1999.

Address reprint requests to Serhan Alkan, M.D., Department of Pathology, Loyola University Medical Center, EMS Building, Suite 2230, 2160 S. First Avenue, Maywood, IL 60153-5385. E-mail: SALKAN@luc.edu.

promoting genes, Ced-3 and Ced-4, which are critical for apoptosis in *C. elegans*.<sup>4</sup> The gene product of Ced-3 demonstrates homology to the mammalian interleukin-1 $\beta$ -converting enzymes (ICE), a group of cysteine proteases.<sup>5</sup> Ced-4 is thought to be homologous to Apaf-1, a mammalian protein that can associate with several death proteases to promote apoptosis.<sup>6</sup>

To date at least 13 members of the ICE/Ced-3 family (caspases) have been identified, the majority of which, on activation, are involved in the induction and execution phases of apoptosis.<sup>7,8</sup> Of these cysteine proteases, caspase-3 (CPP32, Yama, apopain) is believed to be one of the most commonly involved in the execution of apoptosis in various cell types.<sup>7</sup> On cleavage by other caspases, caspase-3 gives rise to two active subunits with molecular masses of 17–20 kd and 10–12 kd.<sup>7,9</sup> These subunits assemble to form an enzymatically active tetrameric complex.<sup>9</sup> Activation of caspase-3 has been described in a number of cell types undergoing apoptosis induced by a variety of stimuli, including CD95 (Fas/Apo-1) signaling.<sup>10,11</sup>

CD95, a cell surface protein receptor belonging to the tumor necrosis factor (TNF)/nerve growth factor receptor family, is an important molecule in the induction of apoptosis in both hematopoietic and nonhematopoietic cells.<sup>9–15</sup> Mutations in the gene that codes for CD95 have been linked to the development of autoimmune disease and lymphoproliferative disorders in both humans and animal models.<sup>16–18</sup> Previous studies demonstrated that crosslinking of the CD95 receptor on the cell surface by agonistic antibody or by its ligand, CD95L, induced apoptosis that was dependent on caspase activation.<sup>19–23</sup> Furthermore, the inhibition of CD95-mediated apoptosis by blocking proteolysis of caspase-3 by viral proteins is suggested to play a role in the pathogenesis of various neoplasms.<sup>10,11,14,15</sup>

The role of caspases, including caspase-3, applied to apoptotic processes in Hodgkin's disease is currently undefined. In this report, we demonstrate caspase-3 plays an important role in CD95-mediated apoptosis in classical Hodgkin's disease (CHD) cell lines. Furthermore, we demonstrate that nodular lymphocyte predominance Hodgkin's disease (NLPHD) lacks caspase-3 expression by immunophenotypic analysis. The lack of caspase-3 expression in NLPHD may contribute to the development and pathogenesis of this disease by imbuing tumor cells with resistance to caspase-3-dependent apoptotic pathways.

## Materials and Methods

### Case Selection, Histological Examination, and Immunohistochemistry of NLPHD

Formalin-fixed, paraffin-embedded tissue sections from 11 cases of NLPHD were selected from the surgical pathology files of Loyola University Medical Center and the University of Michigan Medical School for immunohistochemical determination of caspase-3. Diagnosis of NLPHD was performed using established criteria on

Table 1. Antibodies Used in Immunohistochemical Staining of NLPHD

Antibody	Source	Dilution
LCA (PD7/26/16 and 2B11)	DAKO (Carpinteria, CA)	1:50
CD30 (Ber-H2)	DAKO	1:40
CD20 (L26)	DAKO	1:100
EMA (E29)	DAKO	1:100
CD15 (Leu-M1)	Becton Dickinson, (San Jose, CA)	1:50
CD45RO (A6)	Zymed Laboratories (San Francisco, CA)	1:50

lymph node biopsy histology and immunohistochemistry.<sup>24,25</sup> NLPHD was diagnosed by the finding of typical nodular architecture and lymphocytic/histiocytic (L&H) cells with the appropriate CD20- and CD45RB-positive immunophenotype.

Morphology assessment of NLPHD cases was performed on 4- $\mu$ m tissue sections with hematoxylin-eosin. Immunoperoxidase staining of lymph node sections with the antibodies listed in Table 1 was performed using a Ventana 320 automated stainer (Ventana Medical Systems, Tucson, AZ) and a streptavidin/horseradish peroxidase detection kit (Ventana), with microwave antigen retrieval and trypsin pretreatment used as necessary. The chromogen was 3,3'-diaminobenzidine tetrahydrochloride (DAB).

### Cell Lines

The CHD cell lines KMH2, L428, and HS445 were used in this study. KMH2 and L428 cell lines were obtained from the German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany). HS445 and Jurkat cell lines were obtained from the American Type Culture Collection (Manassas, VA). Cell lines were cultured in RPMI 1640 (Gibco-BRL, Grand Island, NY) supplemented with 20% (v/v) heat-inactivated fetal bovine serum (Sigma Chemical Co., St. Louis, MO), 2 mmol/L L-glutamine (Gibco-BRL), 25 mmol/L Hepes (Sigma), and antibiotic-antimycotic solution (Sigma). All cell lines were maintained at 37°C in a humidified incubator at 5% CO<sub>2</sub>.

### Immunohistochemical Analysis for Caspase-3 Expression

Four-micron-thick formalin-fixed, paraffin-embedded tissue sections from each case of NLPHD were deparaffinized in xylene, hydrated in graded alcohol, and pretreated for antigen retrieval in 10 mmol/L citrate buffer, pH 6.0, for 10 minutes. Cytospins from CHD cell lines were fixed in a 1:1 mixture of acetone and methanol for 10 minutes. Staining was performed using polyclonal rabbit anti-human CPP32 (1:200 titer, DAKO Corp., Carpinteria, CA) and a Vectastain ABC peroxidase, rabbit IgG detection kit (Vector Laboratories, Burlingame, CA) with

3-amino 9-ethyl carbazole (AEC) as the chromogen. The chromogen DAB was used for paraffin-embedded specimens. Formalin-fixed, paraffin-embedded tissue sections from three cases of caspase-3-positive nodular sclerosis Hodgkin's disease and a reactive tonsil were used as positive controls for caspase-3 staining.

### *Apoptosis Induction and Detection*

For apoptosis assays,  $1 \times 10^6$  cells from each cell line were cultured in 24-well tissue culture plates (Falcon, Lincoln Park, NJ) and incubated with 500 ng/ml of agonistic anti-CD95 monoclonal antibody (mAb) (clone CH11, mouse IgM, Upstate Biotechnology, Lake Placid, NY) for indicated time periods, with or without 1 hour preincubation with 10  $\mu$ mol/L caspase-3 peptide inhibitor Ac-Asp-Glu-Val-Asp-fluoromethyl ketone (DEVD-FMK, Clontech, Palo Alto, CA).

Detection of apoptosis in CHD cell lines by terminal deoxynucleotidyl transferase-mediated dUTP nick end-labeling (TUNEL) was quantitated using the ApopTag *in situ* apoptosis peroxidase detection kit (Oncor, Gaithersburg, MD). Cytospin preparations of cells were fixed in 1% formaldehyde for 15 minutes followed by 1 hour fixation in 70% ethanol at  $-20^\circ\text{C}$ . After a brief wash in FA buffer (Difco Laboratories, Detroit, MI), each slide was incubated at room temperature (RT) for 10 minutes with equilibration buffer followed by 1 hour incubation at  $37^\circ\text{C}$  with TdT enzyme (or deionized water ( $\text{dH}_2\text{O}$ ) for negative controls) diluted with the reaction buffer. The TdT reaction was stopped with stop/wash buffer and each specimen was briefly washed with FA buffer before 30 minute incubation with anti-digoxigenin-peroxidase at RT. After a series of washes with FA buffer, each slide was developed with DAB/hydrogen peroxide (Sigma) color substrate for 6 minutes at RT. All slides were counterstained with hematoxylin. A CD95-sensitive Jurkat T cell line was used as a positive control for apoptosis. A positive reaction for apoptosis was characterized by brown/black coloration of the nuclear or perinuclear region of the cell. Apoptotic cells were quantitated by 1000-cell count at  $400\times$  magnification.

The 7-Amino Actinomycin D (7-AAD) staining method to measure cell viability was performed per manufacturer's protocol using Via-Probe 7-AAD (PharMingen, San Diego, CA). Briefly, anti-CD95 mAb-treated and untreated cells ( $1 \times 10^6$  cell/ml) were washed twice in cold PBS and resuspended in  $1\times$  binding buffer (10 mmol/L HEPES/NaOH (pH 7.4), 140 mmol/L NaCl, and 2.5 mmol/L  $\text{CaCl}_2$ ). Resuspended cells were then incubated for 20 minutes at  $20-25^\circ\text{C}$  in the dark with 5  $\mu$ l of 7-AAD. Samples (30,000 events per sample) were then quantitated on an Epics XL-MCL flow cytometer (Coulter, Miami Lakes, FL), recorded in LIST mode, and registered on logarithmic scales. 7-AAD emission was detected in the FL-3 channel ( $>650$  nm). Analysis was performed using Coulter System II software.

### *Determination of Caspase-3 Activity in Cell Lines*

Caspase-3 activity was determined using the ApoAlert CPP32/Caspase-3 colorimetric assay kit (Clontech). After a 48-hour incubation with anti-CD95 mAb, duplicate samples of untreated and treated cells ( $2 \times 10^6$  cells) were washed in cold PBS, resuspended in 50  $\mu$ l cell lysis buffer, and incubated on ice for 10 minutes. Cell lysates were pelleted, followed by transfer of the supernatants to microcentrifuge tubes. Fifty microliters of  $2\times$  reaction buffer with 5 mmol/L DTT and 5  $\mu$ l of 1 mmol/L DEVD-p-nitroanilide (pNA)-conjugated CPP32 substrate were added to each tube, followed by 1 hour incubation in a water bath at  $37^\circ\text{C}$ . A control reaction of treated cells without DEVD-pNA was included. Optical density (OD) for each specimen was determined at 405 nm using the EL 312e microplate reader (Bio-Tek Instruments, Winooski, VT). For quantification of protease activity, sample values were plotted on a calibration curve derived from the OD values obtained from each of five standards (range: 0–20 nmole pNA). For each sample, units of CPP32 activity were determined by the following formula:

$$\text{Units of CPP32 activity} = (\Delta\text{OD})$$

$$\times (\text{calib. Curve slope})^{-1} \times (\text{nmole pNA/OD})$$

where  $\Delta\text{OD}$  is the change in optical density from the control reaction without conjugated substrate.

## *Results*

### *Histology and Immunohistochemical Characterization of NLPHD Cases*

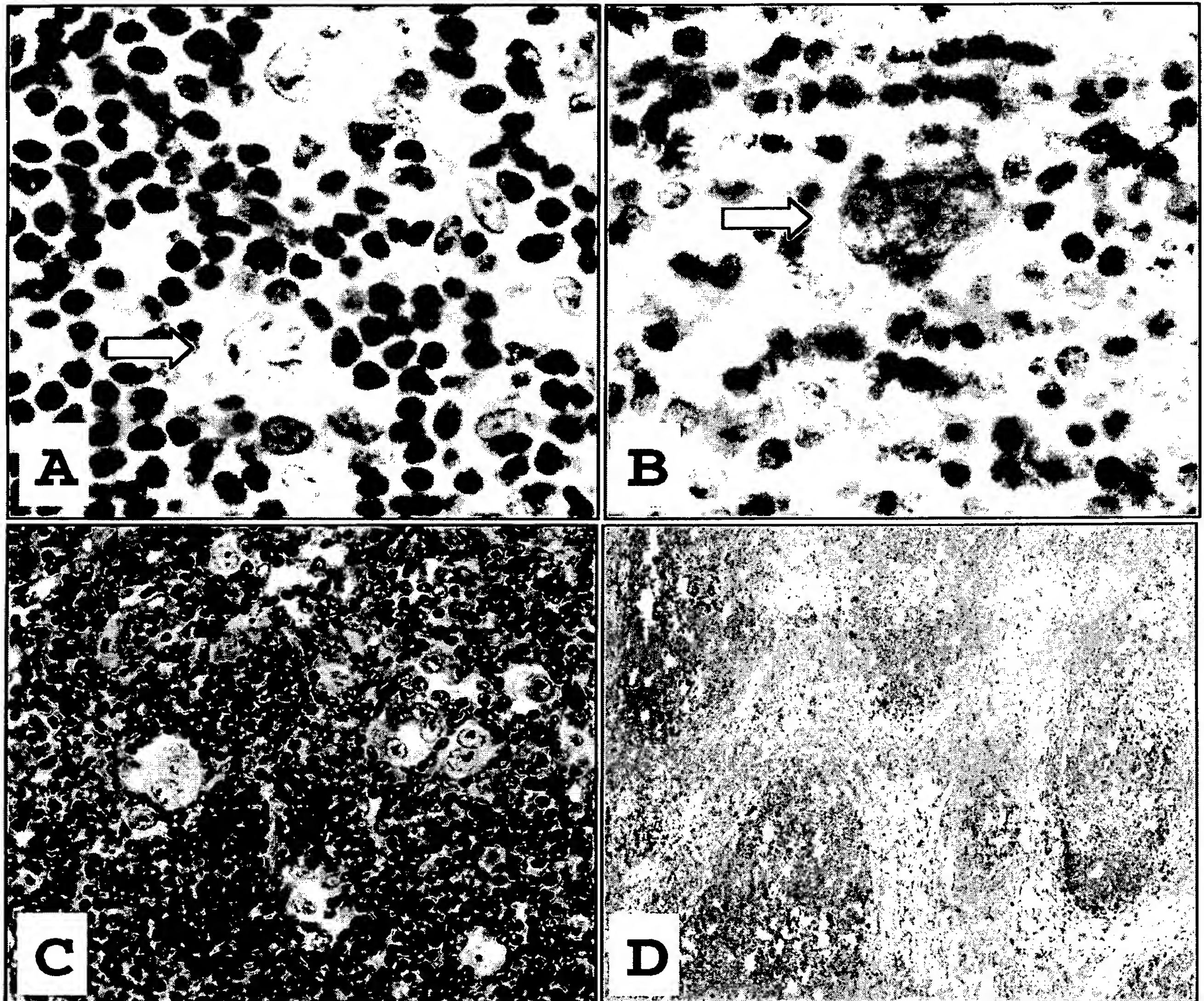
All 11 specimens demonstrated architectural changes consistent with NLPHD and were composed predominantly of large nodules with focal areas of diffuse effacement. Typical nodules contained characteristic L&H cells intermixed in a background of small lymphocytes and occasional epithelioid histiocytes separated by compressed intervening paracortical areas composed of small lymphocytes and scattered plasma cells.

In all specimens, L&H cells demonstrated positive staining for CD20 and LCA. In seven specimens, L&H cells expressed EMA, with one case also expressing CD30. In all cases, L&H cells were negative for CD15 and CD45RO.

### *Most NLPHD Cases Failed to Express Caspase-3 by Immunohistochemistry*

In 10 of 11 cases of NLPHD, including the case which expressed CD30, L&H cells were negative for caspase-3 expression by immunohistochemical staining as represented in Figure 1A. In one case, caspase-3 immunopositivity was detected in the cytoplasm in a few scattered L&H cells (Figure 1B). In contrast, caspase-3 expression was demonstrated both in Hodgkin Reed-Sternberg





**Figure 1.** This L&H cell (arrow), as seen in the majority of NLPHD cases, was immunohistochemically negative for caspase-3, whereas scattered plasma cells and lymphocytes expressed caspase-3 (A, DAB  $\times 1000$ ). A single L&H cell (arrow) from one case of NLPHD displayed cytoplasmic expression of caspase-3 (B, DAB  $\times 1000$ ). Control cases of nodular sclerosis Hodgkin's disease demonstrated diffuse caspase-3-immunopositivity of HRS cells and intense positive immunostaining of lymphocytes and plasma cells within the surrounding infiltrate (C, DAB  $\times 400$ ). Reactive follicular centers in tonsil controls also displayed intense positive staining for caspase-3 (D, DAB,  $\times 200$ ).

(HRS) cells and in background lymphocytes in three cases of nodular sclerosis Hodgkin's disease (Figure 1C). In addition, tonsil tissue positive controls demonstrated caspase-3 immunopositivity concentrated predominantly in germinal center cells of secondary follicles (Figure 1D).

#### *Caspase-3 Was Detectable in CHD Cell Lines*

Three CHD cell lines (HS445, L428, and KMH2) were analyzed for caspase-3 expression by immunohistochemistry. HS445 and L428 consistently demonstrated substantial cytoplasmic immunostaining for caspase-3 (Figure 2 and data not shown). However, in contrast,

repeated immunohistochemistry assays failed to detect expression of caspase-3 in the KMH2 cell line (Figure 2).

#### *Caspase-3 Was Proteolytically Cleaved and Activated during CD95-Mediated Apoptosis in Caspase-3-Positive CHD Cell Lines*

Activation of the CD95 receptor by ligand or agonistic mAb is known to induce apoptosis with concomitant proteolytic cleavage and activation of caspases, including caspase-3, in CD95-positive neoplasms.<sup>19-23</sup> To investigate the effect of CD95 stimulation with potential activation of caspase-3 in Hodgkin's disease, we examined the



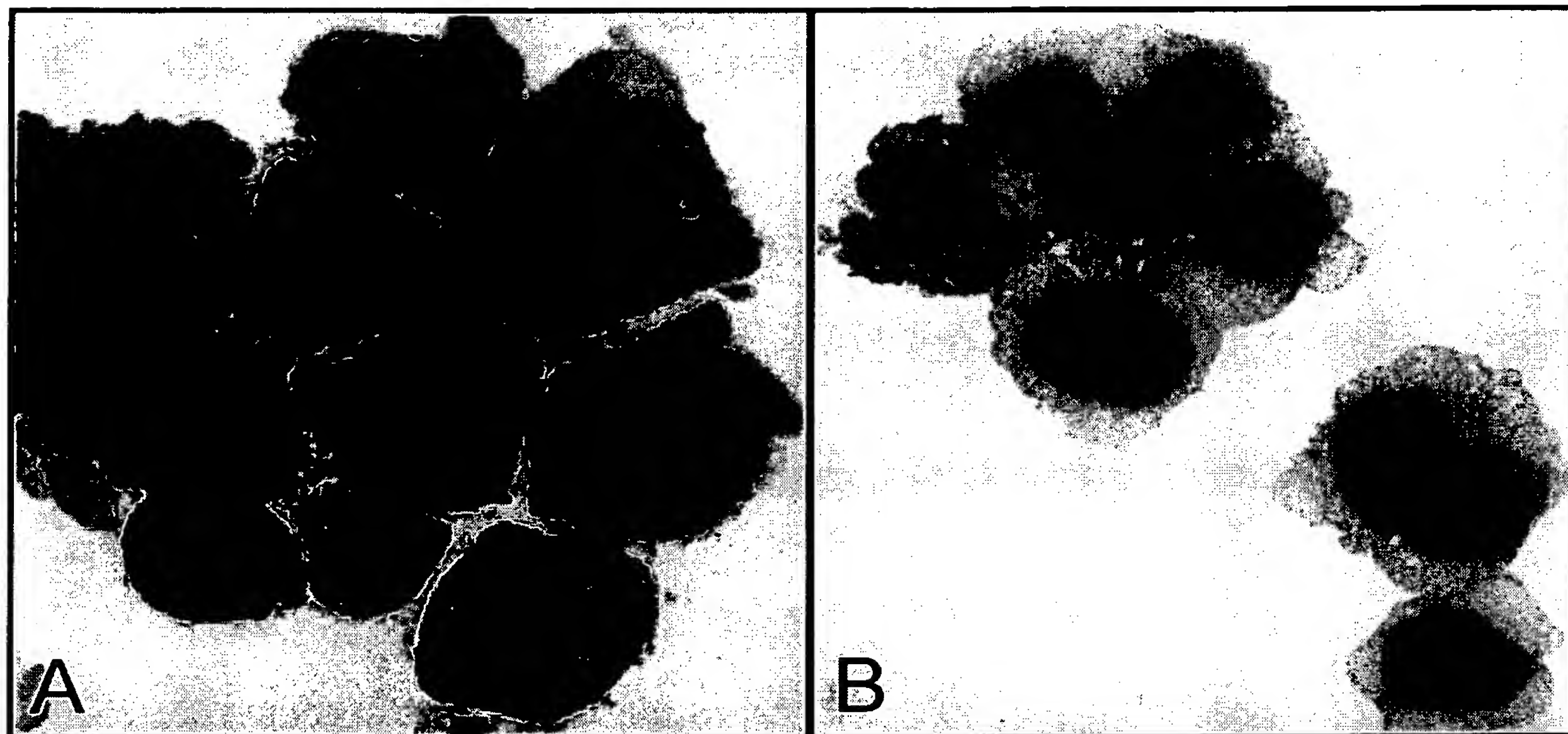


Figure 2. Immunohistochemical detection of caspase-3 on cytospin preparations of the L428 cell line (A) displayed strong cytoplasmic positive staining for caspase-3 (AEC,  $\times 400$ ); however, the KMH2 cell line (B) failed to express caspase-3 (AEC,  $\times 400$ ). Isotype control antibody staining was negative (data not shown).

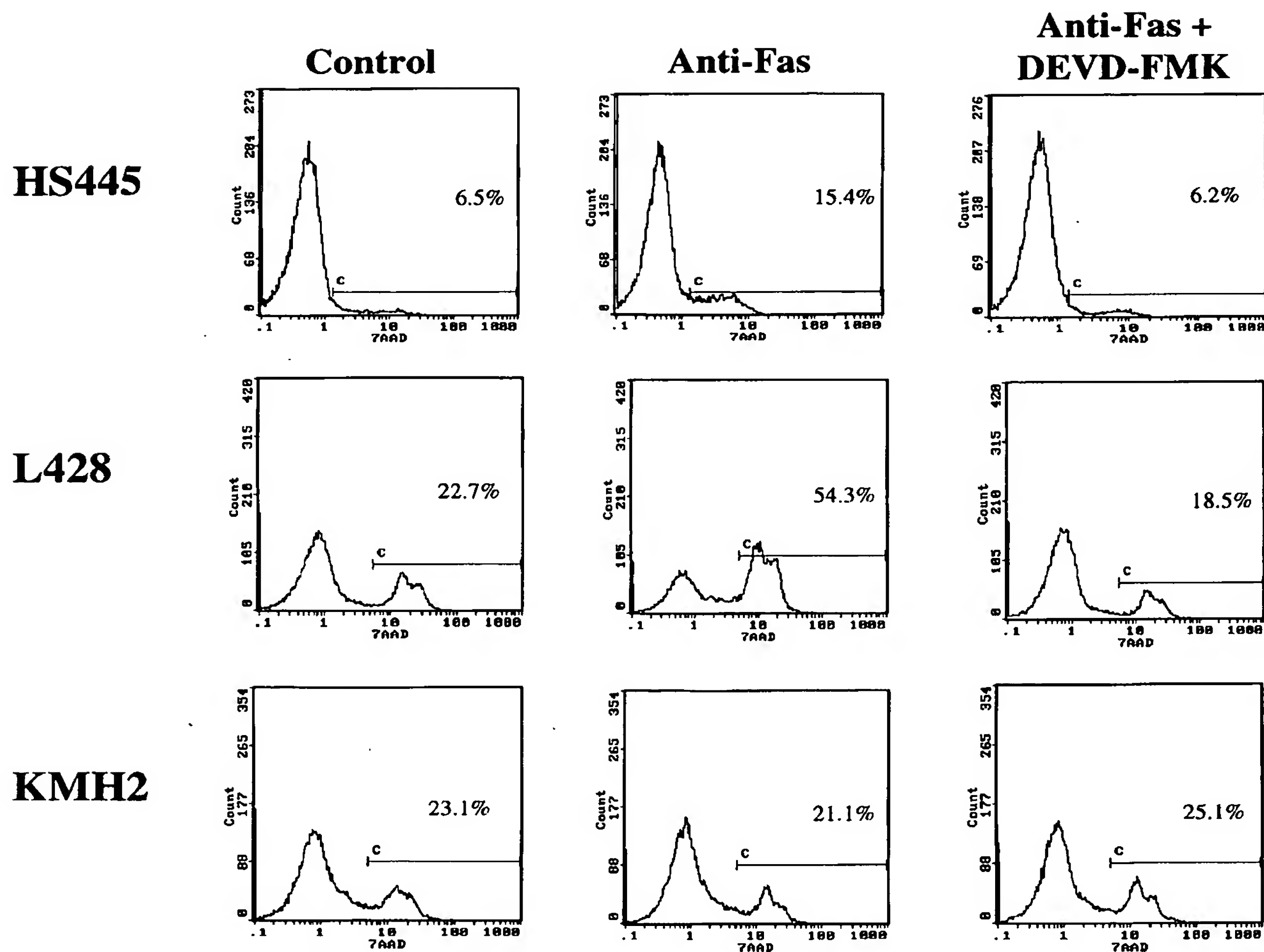
effect of agonistic CD95 mAb on CHD cell lines. The HS445 and L428 cell lines displayed a significant increase in apoptosis after 72 hours' treatment with anti-CD95 mAb as quantitated by both flow cytometric analysis with 7-AAD (Figure 3) staining and the TUNEL assay (Figure 4, Table 2). CD95-induced apoptosis was significantly inhibited in these cells by the caspase-3 peptide inhibitor, DEVD-FMK, as demonstrated by the decrease in number of apoptotic cells to nearly background levels (Figures 3 and 4, Table 2). In contrast, the KMH2 cell line demonstrated consistent resistance to anti-CD95 mAb treatment with no effect by the addition of DEVD-FMK (Figure 3 and Table 2). Resistance of KMH2 cells to CD95-mediated apoptosis was not due to lack of CD95 expression, as all three cell lines expressed similarly high levels of CD95 as determined by flow cytometric staining.<sup>26</sup>

The significant inhibition of apoptosis by a caspase-3 inhibitor in HS445 and L428 cells, and the lack of apoptosis induced by caspase-3-deficient KMH2 cells suggests caspase-3 is important for CD95-mediated apoptosis in CHD cell lines. To further substantiate caspase-3 cleavage and activation in CD95-mediated apoptosis in CHD, each cell line was evaluated for changes in caspase-3 activity before and after treatment with anti-CD95 mAb using an enzyme colorimetric assay (Figure 5). Forty-eight-hour treatment revealed approximately tenfold increases in caspase-3 activity in HS445 and L428 cells in contrast to no difference detected in treated KMH2 cells. Positive control CD95-sensitive Jurkat T cells displayed a fivefold increase in caspase-3 activity after 24 hours of treatment with anti-CD95 mAb.

## Discussion

Among the caspases identified in humans thus far, caspase-3 is probably one of the most relevant and best studied as regards to apoptosis in hematopoietic cells. Caspase-3 (CPP32, Yama, apopain) has been shown to be a key effector molecule in the downstream execution of various apoptotic stimuli.<sup>9-11,27-30</sup> Activated caspase-3 cleaves and inactivates many vital cellular proteins during apoptosis including kinases and proteins associated with cellular structure, cell cycle, and DNA repair. One such well characterized caspase-3 death substrate is poly(ADP-ribose) polymerase (PARP), an enzyme involved in DNA repair, genome surveillance, and integrity.<sup>31,32</sup> In addition, caspase-3 appears to indirectly activate endonucleases implicated in internucleosomal DNA cleavage by removing the negative regulatory effect of PARP.<sup>31</sup>

The cleavage and activation of caspase-3 during apoptosis has been well documented in neoplastic cells. Caspase-3 activation and subsequent cleavage of its substrates, protein kinase C- $\delta$  (PKC- $\delta$ ) and PARP, was demonstrated by chemotherapeutic drug treatment in human leukemic cell lines.<sup>33</sup> Cross-resistance to CD95- and chemotherapeutic drug-induced apoptosis due to lack of caspase activation including caspase-3 was demonstrated in a human acute T-cell leukemia line, CEM.<sup>34</sup> Also, MCF breast carcinoma cells lacking expression of caspase-3 were resistant to apoptotic stimuli.<sup>35</sup> Thus, the expression and activation of caspase-3 appears to be critical for the execution of various apoptotic stimuli in neoplasms.



**Figure 3.** Flow cytometric analysis of 7-AAD staining in anti-CD95 (Fas)-treated CHD cell lines. Increased cell death was observed after 72 hours of treatment with 500 ng/ml agonistic CD95 mAb (CH11) in HS445 and L428 cell lines compared to untreated control cells. Pretreatment of cells with caspase-3 peptide inhibitor, DEVD-FMK, significantly decreased cell death in anti-CD95-treated cells to near background levels. In contrast, no significant increase in cell death was observed after anti-CD95 treatment of KMH2 cells as compared to untreated cells. The x axis represents fluorescence intensity (log scale) and the y axis represents relative cell number. These data are representative of at least three separate experiments performed.

Hodgkin's disease accounts for 14% of malignant lymphomas. Currently, one-third of advanced Hodgkin's disease patients are resistant to conventional therapies.<sup>36</sup> Our knowledge of the expression and function of apoptosis-related proteins such as caspases and how they may contribute to the pathogenesis and treatment of this malignancy is limited. Previous immunohistochemical studies *in situ* demonstrated that caspase-3 is commonly expressed in CHD.<sup>37</sup> However, the examination of caspase-3 expression in NLPHD has been limited. Furthermore, the overall biological significance of caspase-3 in Hodgkin's disease is unknown. Therefore, in this study we examined the expression of caspase-3 in NLPHD and determined its functional significance in CHD cell lines.

We first examined the *in situ* expression of caspase-3 in NLPHD. By immunohistochemistry, we identified caspase-3 immunopositivity in scattered L&H cells from only 1 of 11 cases of NLPHD. These findings are in

agreement with the study of Chhanabhai and colleagues, who found no expression of caspase-3 in L&H cells from 6 cases of NLPHD.<sup>37</sup> In addition, these authors observed the HRS in the majority of cases of CHD were positive for caspase-3 expression.<sup>37</sup> These latter observations differ from the immunohistochemical findings of Xerri et al in which only 3 of 16 cases of HRS of CHD (nodular sclerosis and mixed cellularity type) were caspase-3-immunopositive.<sup>38</sup> The reason for the difference in caspase-3 expression in CHD noted between these groups is presently unclear.

Our immunohistochemical analysis of HRS cells of three CHD cell lines revealed substantial expression of caspase-3 in HS445 and L428, but only weak expression in KMH2 cells. These findings concur with Western blot analysis of these cell lines, which revealed expression of the 32-kd zymogen form of caspase-3 in HS445 and L428, but virtually undetectable expression in KMH2 cells.<sup>26</sup>

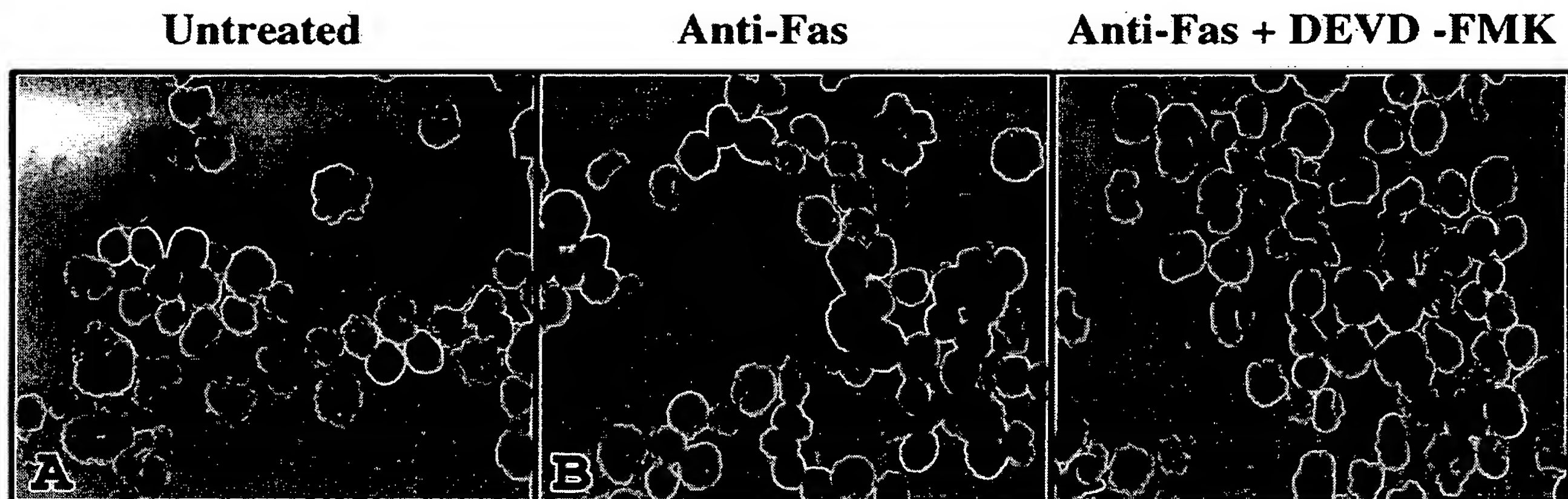


Figure 4. TUNEL assay for apoptosis. Compared to untreated cells (A), the L428 cell line displayed a considerable increase in apoptosis (dark brown/black cells) after 72 hour incubation with 500 ng/ml agonistic CD95 (Fas) mAb (B). Pretreatment of L428 cells with 10  $\mu$ M DEVD-FMK decreased the number of apoptotic cells to near baseline levels (C). Stimulation of KMH2 cells with anti-CD95 mAb with or without DEVD-FMK pretreatment showed no increase in apoptosis compared to untreated cells (data not shown).

To address the biological significance of caspase-3 in Hodgkin's disease, we investigated the role of caspase-3 in CD95-mediated apoptosis in CHD lines. After stimulation of the CD95 receptor by agonistic CD95 mAb, significant apoptosis was induced in caspase-3 positive cell lines HS445 and L428 by TUNEL and 7-AAD assays. However, KMH2 cells, which virtually failed to express caspase-3 by immunohistochemistry, were consistently

resistant to CD95 stimulation, suggesting that resistance to CD95-mediated apoptosis in this cell line may be due to a deficiency of caspase-3<sup>26</sup> (manuscript in preparation).

To establish caspase-3 as a key mediator in CD95-induced apoptosis in CHD cell lines, enzyme assays specific for caspase-3 activity were performed. Approximately tenfold increases in caspase-3 activity were observed in HS445 and L428 after 48 hours' incubation with anti-CD95 mAb, compared to no increase in KMH2 cells. In addition, we pretreated each cell line with the caspase-3 peptide inhibitor DEVD-FMK before CD95 activation. Previous studies in other experimental systems have demonstrated that DEVD inhibitors have specificity for caspase-3 by bearing similarities to the cleavage site of the caspase-3 substrate, PARP.<sup>9,11,27-31,39-41</sup> The addition of DEVD-FMK to cultures of HS445 and L428 significantly decreased CD95-mediated apoptosis; however, there was no effect on KMH2 cells. These findings in CHD lines correlate with previous studies which demonstrated caspase-3 is proteolytically cleaved and activated and plays a key role in CD95-mediated apoptosis in other experimental systems.<sup>9-11,27-30,42</sup> However, it should be noted that CD95-induced apoptosis may occur without activation of caspase-3, suggesting the existence of alternate apoptosis execution pathways in response to CD95 signaling.<sup>43</sup>

Most investigations related to apoptosis in Hodgkin's disease have focused on the expression of mitochondrial apoptosis regulatory proteins Bcl-2, Bcl-x, and Bax.<sup>44-49</sup> These studies revealed variable expression of Bcl-2<sup>44-48</sup> but frequent expression of the pro-apoptotic protein Bax<sup>47</sup> and the apoptosis antagonist Bcl-x<sub>L</sub>.<sup>48-49</sup> Previous investigations of CD95 expression by HRS cells have been limited; however, these studies revealed that CD95 is expressed on HRS cells in the majority of cases of CHD.<sup>50-53</sup> In this report, we demonstrate that CHD cell lines expressing CD95 can undergo apoptosis by CD95 stimulation. A recent study assayed CD95-induced apoptosis in fresh tissue samples with Hodgkin's disease; however, the HRS cells were not specifically analyzed.<sup>51</sup>

Table 2. Apoptosis Rates Induced in Anti-CD95 mAb Treated or Untreated CHD Cell Lines as Determined by TUNEL Assay

CHD cell line	Apoptosis*		
	Untreated	Anti-CD95	Anti-CD95+ DEVD-FMK
HS445	7.5%	20.5%	9.9%
L428	7.8%	30.2%	10.8%
KMH2	1.0%	1.3%	1.0%

\*1  $\times 10^6$  cells were untreated or treated with 500 ng/ml anti-CD95 mAb (CH11) for 72 hours with or without pretreatment with 10  $\mu$ M caspase-3 peptide inhibitor, DEVD-FMK. Apoptosis was quantitated by TUNEL staining as described in Materials and Methods.

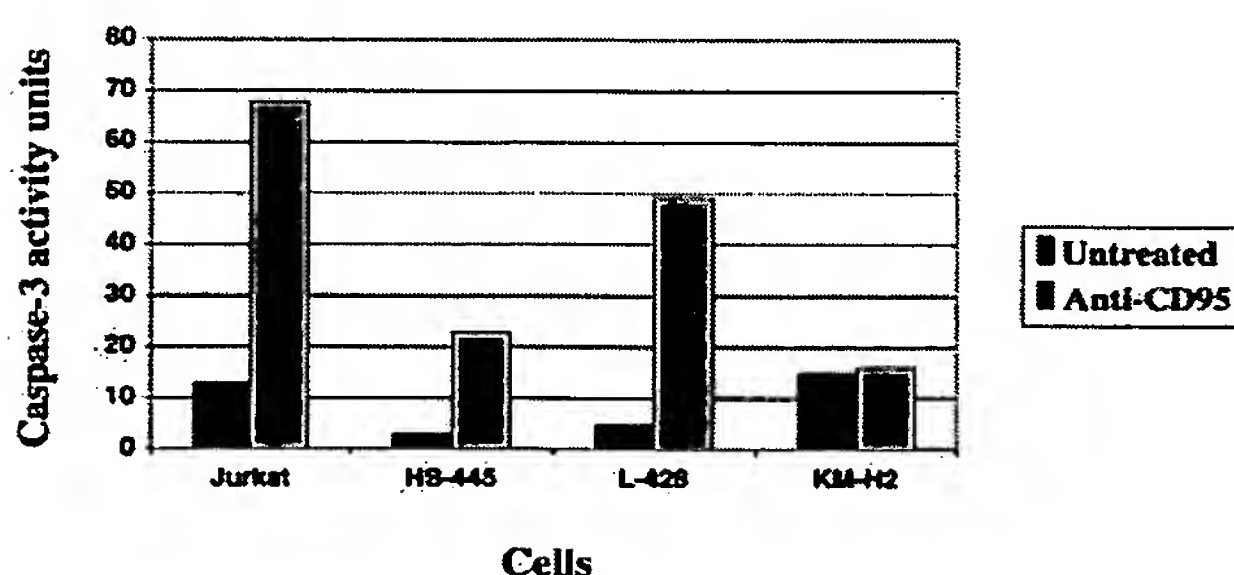


Figure 5. Caspase-3 activity in anti-CD95-treated CHD cell lines. Positive control for caspase-3 activity is demonstrated in Jurkat T cells by fivefold induction of caspase-3 activity after 24 hours of agonistic CD95 mAb treatment. HS445 and L428 cell lines had approximately tenfold increases in caspase-3 activity after 48 hours' treatment with anti-CD95 mAb compared to untreated controls. In contrast, no significant increase in caspase-3 activity was detected in anti-CD95-treated KMH2 cells when compared to untreated cells. These data are representative of three separate experiments.



Because HRS cells usually constitute less than 1% of involved tissue, it is difficult to assay CD95 stimulation of HRS cells directly without selective separation.

The absence of caspase-3 expression in L&H cells is similar to that seen in several indolent B-cell NHLs, most notably follicular center lymphoma (FCL), grade I.<sup>38,54</sup> Recent studies have noted clinical similarities between NLPD and indolent B-cell NHL.<sup>55-57</sup> However, NLPD differs from the majority of low-grade B-cell NHLs with respect to treatment response. Most NLPD patients are cured and rarely show progressive disease, in contrast in the majority of FCL patients. Furthermore, although bcl-2 is commonly overexpressed in FCLs, NLPD typically lacks expression of this protein.<sup>24,46,58</sup> The combined high and low expression of bcl-2 and caspase-3 protein, respectively, in low-grade FCL suggests the incurability of many of these lymphomas may be directly related to the overexpression of anti-apoptotic proteins (eg, Bcl-2) combined with the lack of downstream apoptotic mediators such as caspase-3. Furthermore, overexpression of Bcl-2 and Bcl-x<sub>L</sub> in cell lines can also block cleavage and activation of caspase-3.<sup>27,29,59-61</sup>

The lack of caspase-3 expression in NLPD may also be an important mechanism of resistance to apoptosis. Furthermore, the differential expression of caspase-3 between CHD and NLPD suggests that each may be a distinct disease entity, and may account for some of the clinical differences between these two disorders. Additional studies to define the expression and function of caspases and their relationship to other apoptosis-related proteins may provide novel insights into the pathogenesis and treatment resistance of this malignancy.

## Acknowledgments

We thank Barbara Rozhon for assistance in preparing the manuscript, Luann Desautel of the Loyola University Medical Center Flow Cytometry Laboratory for technical assistance with 7-AAD staining, and Heide Guzlas for expert photographic assistance. The authors gratefully acknowledge Dr. Tom Ellis for critical review of this manuscript.

## References

1. Tsujimoto Y, Croce CM: Analysis of the structure, transcripts, and protein products of bcl-2, the gene involved in human follicular lymphoma. *Proc Natl Acad Sci USA* 1986, 83:5214-5218
2. Vaul DL, Cory S, Adams JM: Bcl-2 gene promotes haematopoietic cell survival and cooperates with c-myc to immortalize pre-B cells. *Nature* 1988, 335:440-442
3. Hengartner MO, Horvitz HR: *C. elegans* survival gene ced-9 encodes a functional homologue of the mammalian proto-oncogene bcl-2. *Cell* 1994, 76:665-676
4. Shaham S, Horvitz HR: Developing *Caenorhabditis elegans* neurons may contain both cell-death protective and killer activities. *Genes Dev* 1996, 10:578-591
5. Yuan J, Shaham S, Ledoux S, Ellis JM, Horvitz JR: *C. elegans* cell death gene ced-3 encodes a protein similar to mammalian interleukin-1 $\beta$ -converting enzyme. *Cell* 1993, 75:641-652
6. Hu Y, Benedict MA, Wu D, Inohara N, Nunez G: Bcl-XL interacts with Apaf-1, and inhibits Apaf-1-dependent caspase-9 activation. *Proc Natl Acad Sci USA* 1998, 95:4386-4391
7. Alnemri ES, Livingston DJ, Nicholson DW, Salvesen G, Thornberry NA, Wong WW, Yuan J: Human ICE/CED-3 protease nomenclature. *Cell* 1996, 87:171
8. Humke EW, Ni J, Dixit VM: ERICE, a novel FLICE-activatable caspase. *J Biol Chem* 1998, 273:15702-15707
9. Schlegel J, Peters I, Orrenius S, Miller DK, Thornberry NA, Yamin TT, Nicholson DW: CPP32/apopain is a key interleukin 1 $\beta$  converting enzyme-like protease involved in Fas-mediated apoptosis. *J Biol Chem* 1996, 1841-1844
10. Enari M, Hug H, Nagata S: Involvement of an ICE-like protease in Fas-mediated apoptosis. *Nature* 1995, 375:78-81
11. Enari M, Talianian RV, Wong WW, Nagata S: Sequential activation of ICE-like and CPP32-like proteases during Fas-mediated apoptosis. *Nature* 1996, 380:723-726
12. Itoh N, Yonehara S, Ishii A, Yonehara M, Mizushima S, Sameshima M, Hase A, Seto Y, Nagata S: The polypeptide encoded by the cDNA for human cell surface antigen Fas can mediate apoptosis. *Cell* 1991, 66:233-243
13. Oehm A, Behrmann I, Falk W, Pawlita M, Maier G, Klas C, Li-Weber M, Richards S, Dhein J, Trauth BC: Purification and molecular cloning of the APO-1 cell surface antigen, a member of the tumor necrosis factor/nerve growth factor receptor superfamily. Sequence identity with the Fas antigen. *J Biol Chem* 1992, 267:10709-10715
14. Smith CA, Farrah T, Goodwin RG: The TNF receptor superfamily of cellular, and viral proteins: activation, costimulation and death. *Cell* 1994, 76:959-962
15. Bump NJ, Hackett M, Hugunin M, Seshagiri S, Brady K, Chen P, Ferenz C, Franklin S, Ghayur T, Li P: Inhibition of ICE family proteases by baculovirus antiapoptotic protein p35. *Science* 1995, 269:1885-1888
16. Nagata S, Suda T: Fas and Fas ligand: lpr and gld mutations. *Immunol Today* 1995, 16:39-43
17. Takahashi T, Tanaka M, Brannan CI, Jenkins NA, Copeland NG, Suda T, Nagata S: Generalized lymphoproliferative disease in mice caused by a point mutation in Fas ligand. *Cell* 1994, 76:969-976
18. Gronbaek K, Thor Straten P, Ahrenkeil V, Klarskov Anderson M, Ebbe Hansen N, Zeuthen J, Hou-Jensen K, Guldberg P: Somatic fas mutations in non-Hodgkin's lymphomas: association with extranodal disease and autoimmunity. *Blood* 1998, 92:3018-3024
19. Trauth BC, Klas C, Peters AM, Matzku S, Moller P, Falk W, Debatin KM, Krammer PH: Monoclonal antibody-mediated tumor regression by induction of apoptosis. *Science* 1989, 245:301-305
20. Yonehara S, Ishii A, Yonehara M: A cell killing monoclonal antibody (anti-Fas) to a cell surface antigen co-downregulated with the receptor of tumor necrosis factor. *J Exp Med* 1989, 169:1747-1756
21. Suda T, Nagata S: Purification and characterization of the Fas ligand that induces apoptosis. *J Exp Med* 1994, 179:873-878
22. Suda T, Takahashi T, Golstein P, Nagata S: Molecular cloning and expression of the Fas ligand, a novel member of the TNF family. *Cell* 1993, 75:1169-1178
23. Takahashi T, Tanaka M, Inazawa J, Abe T, Suda T, Nagata S: Human Fas ligand: gene structure, chromosomal location and species specificity. *Int Immunol* 1994, 6:1567-1574
24. von Wasielewski R, Werner M, Fischer R, Hansmann ML, Hubner K, Hasenclever D, Franklin J, Sextro M, Diehl V, Georgii A: Lymphocyte-predominant Hodgkin's disease: An immunohistochemical analysis of 208 reviewed Hodgkin's disease cases from the German Hodgkin's Study Group. *Am J Pathol* 1997, 150:793-803
25. Harris NL, Jaffe ES, Stein H, Banks P, Chan JK, Cleary ML, Delsol G, De Wolf-Peters C, Falini B, Gatter KC, Isaacson PG, Knowles DM, Mason DY, Muller-Hermelink H-K, Pileri SA, Piris MA, Ralfkiaer E, Warnke RA: A revised European-American classification of lymphoid neoplasms: A proposal from the International Lymphoma Study Group. *Blood* 1994, 84:1361-1392
26. Alkan S, Hsi ED, Wrono-Smith T: CD95 (Fas antigen: APO-1) mediated apoptosis on Hodgkin's disease cell lines. *Mod Pathol* 1998, 11:162A (abstr. 950)
27. Armstrong RC, Aja T, Xiang J, Gaur S, Krebs JF, Hoang K, Bai X, Korsmeyer SJ, Karanewski DS, Fritz LC, Tomaselli KJ: Fas-induced activation of the cell death-related protease CPP32 is inhibited by Bcl-2 and by ICE family protease inhibitors. *J Biol Chem* 1996, 271:16850-16855
28. Los M, Van de Craen M, Penning LC, Schenk H, Westendorp M, Baeuerle PA, Droge W, Krammer PH, Fiers W, Schulze-Osthoff K:

- Requirement of an ICE/CED-3 protease for Fas/APO-1-mediated apoptosis. *Nature* 1995, 375:81-83
29. Chinnaiyan AM, Orth K, O'Rourke K, Duan H, Poirier GG, Dixit VM: Molecular ordering of the cell death pathway: Bcl-2 and Bcl-x<sub>L</sub> function upstream of the CED-3-like apoptotic proteases. *J Biol Chem* 1996, 271:4573-4576
30. Darmon AJ, Bleackley RC: An interleukin-1 $\beta$  converting enzyme-like protease is a key component of Fas-mediated apoptosis. *J Biol Chem* 1996, 271:21699-21702
31. Nicholson DW, Ali A, Thornberry NA, Vaillancourt JP, Ding CK, Gallant M, Gareau Y, Griffin PR, Labelle M, Lazebnik YA: Identification and inhibition of the ICE/CED-3 protease necessary for mammalian apoptosis. *Nature* 1995, 376:37-43
32. Tewari M, Quan LT, O'Rourke K, Desnoyers S, Zeng Z, Beidler DR, Poirier GG, Salvesen GS, Dixit VM: Yama/CPP32b, a mammalian homologue of CED-3, is a CrmA-inhibitable protease that cleaves the death substrate poly (ADP-ribose) polymerase. *Cell* 1995, 81:801-809
33. Datta R, Banach D, Kojima H, Talanian RV, Alnemri ES, Wong WW, Kufe DW: Activation of the CPP32 protease in apoptosis induced by 1- $\beta$ -D-Arabinofuranosylcytosine and other DNA-damaging agents. *Blood* 1996, 6:1936-1943
34. Los M, Herr I, Friesen C, Fulda S, Schulze-Osthoff K, Debatin K-M: Cross-resistance of CD95- and drug-induced apoptosis as a consequence of deficient activation of caspases (ICE/Ced-3 proteases). *Blood* 1997, 8:3118-3129
35. Li F, Srinivasan A, Wang Y, Armstrong RC, Tomaselli KJ, Fritz LC: Cell-specific induction of apoptosis by microinjection of cytochrome c. *J Biol Chem* 1997, 272:30299-30305
36. Hasenclaver D, and Diehl V: A prognostic score for advanced Hodgkin's disease. *New Engl J Med* 1998, 339:1506-1514
37. Chhanabhai M, Krajewski S, Krajewska M, Wang HG, Reed JC, Gascoyne RD: Immunohistochemical analysis of interleukin-1 $\beta$ -converting enzyme/Ced-3 family protease CPP32/Yama/Caspase 3 in Hodgkin's disease. *Blood* 1997, 90:2451-2455
38. Xerri L, Devillard E, Ayello C, Brousset P, Reed JC, Emile J-F, Hassoun J, Parmentier S, Birg F: Cysteine protease CPP32, but not Ich1-L, is expressed in germinal center B cells and their neoplastic counterparts. *Hum Pathol* 1998, 29:912-921
39. Livingston DJ: In vitro and in vivo studies of ICE inhibitors. *J Cell Biochem* 1997, 64:19-26
40. Shimizu S, Eguchi Y, Kamiike W, Matsuda H, Tsujimoto Y: Bcl-2 expression prevents activation of the ICE protease cascade. *Oncogene* 1996, 12:2251-2257
41. Talanian RV, Quinlan C, Trautz S, Hackett MC, Mankovich JA, Banach D, Ghayur T, Brady KD, Wong WW: Substrate specificities of caspase family proteases. *J Biol Chem* 1997, 272:9677-9682
42. Zhivotovsky B, Burgess DH, Schlegel J, Porn MI, Vanags D, Orrenius S: Proteases in Fas-mediated apoptosis. *J Cell Biochem* 1997, 64:43-49
43. Monney L, Otter I, Olivier R, Ozer HL, Hass AL, Omura S, Borner C: Defects in the ubiquitin pathway induce caspase-independent apoptosis blocked by Bcl-2. *J Biol Chem* 1998, 273:6121-6131
44. Gupta RK, Lister TA, Bodmer JG: The t(14;18) chromosomal translocation and Bcl-2 protein expression in Hodgkin's disease. *Leukemia* 1994, 8:1337-1341
45. Bhagat SK, Medeiros LJ, Weiss LM, Wang J, Raffeld M, Stetler-Stevenson M: Bcl-2 expression in Hodgkin's disease. Correlation with the t(14;18) translocation and Epstein-Barr virus. *Am J Clin Pathol* 1993, 99:604-608
46. Alkan S, Ross CW, Hanson CA, Schnitzer B: Epstein-Barr virus bcl-2 protein overexpression are not detected in the neoplastic cells of nodular lymphocyte predominance Hodgkin's disease. *Mod Pathol* 1995, 8:544-547
47. Rigal-Huguet F, Gopas J, Prinsloo I, Pris J, Delsol G, Reed JC, Schlaifer D, Brousset P, Benharroch D, Krajewski S, Laurent G, Meggetto F: Frequent expression of the cell death-inducing gene Bax in Reed-Sternberg cells of Hodgkin's disease. *Blood* 1996, 87:2470-2475
48. Schlaifer D, March M, Krajewski S, Laurent G, Pris J, Delsol G, Reed JC, Brousset P: High expression of the bcl-x gene in Reed-Sternberg cells of Hodgkin's disease. *Blood* 1995, 85:2671-2674
49. Schlaifer D, Krajewski S, Rigal-Huguet F, Laurent G, Pris J, Delsol G, Reed JC, Brousset P: Bcl-x gene expression in Hodgkin's disease. *Leuk Lymphoma* 1996, 23:143-146
50. Kubonishi I, Daibata M, Sakuma I, Yoshino T, Sonobe H, Ohtsuki Y, Miyoshi I: Expression of Fas and apoptosis of a Hodgkin's disease cell line (HD-70). *Int J Hematol* 1997, 65:305-307
51. Xerri L, Devillard E, Hassoun J, Haddad P, Birg F: Malignant and reactive cells from human lymphomas frequently express Fas ligand but display a different sensitivity to Fas-mediated apoptosis. *Leukemia* 1997, 11:1868-1877
52. Nguyen PL, Harris NL, Ritz J, Robertson MJ: Expression of CD95 antigen and Bcl-2 protein in non-Hodgkin's lymphomas and Hodgkin's disease. *Am J Pathol* 1996, 148:847-853
53. Dirks W, Schone S, Uphoff C, Quentmeier H, Pradella S, Drexler HG: Expression and function of CD95 (FAS/APO-1) in leukaemia-lymphoma tumour lines. *Br J Haematol* 1997, 96:584-593
54. Krajewski S, Gascoyne RD, Zapata JM, Krajewska M, Kitada S, Chhanabhai M, Horsmann D, Berean K, Piro LD, Fugier-Vivier I, Liu YJ, Wang HG, Reed JC: Immunolocalization of the ICE/Ced-3-family protease, CPP32 (Caspase 3), in non-Hodgkin's lymphomas, chronic lymphocytic leukemias, and reactive lymph nodes. *Blood* 1997, 89:3817-3825
55. Pinkus GS, Said JW: Hodgkin's disease, lymphocyte predominance type, nodular- further evidence for a B-cell derivation. *Am J Pathol* 1988, 133:211-217
56. Grossman DM, Hanson CA, Schnitzer B: Simultaneous lymphocyte predominant Hodgkin's disease and large-cell lymphoma. *Am J Surg Pathol* 1991, 15:668-676
57. Sextro M, Diehl V, Franklin J, Hansmann ML, Anagnostopoulous I, Marafioti T, Stein H: Lymphocyte predominant Hodgkin's disease: a workshop report. *Ann Oncol* 1996, 7:61-65
58. Algara P, Martinez P, Sanchez L, Villuendas R, Orradre JL, Oliva H, Piris MA: Lymphocyte predominance Hodgkin's disease (nodular paraganuloma): a bcl-2 negative germinal centre lymphoma. *Histopathology* 1991, 19:69-75
59. Ibrado AM, Huang Y, Fang G, Liu L, Bhalla K: Overexpression of Bcl-2 or Bcl-x<sub>L</sub> inhibits Ara-C-induced CPP32/Yama protease activity and apoptosis of human acute myelogenous leukemia HL-60 cells. *Cancer Res* 1996, 56:4743-4748
60. Memon SA, Moreno MB, Petrak D, Zacharchuk CM: Bcl-2 blocks glucocorticoid- but not Fas- or activation-induced apoptosis in a T cell hybridoma. *J Immunol* 1995, 155:4644-4652
61. Moreno MB, Memon SA, Zacharchuk CM: Apoptosis signaling pathways in normal T cells: differential activity of Bcl-2 and IL-1 $\beta$ -converting enzyme family protease inhibitors on glucocorticoid- and Fas-mediated cytotoxicity. *J Immunol* 1996, 157:3845-3849

## Purification and characterization of a human tumor necrosis factor from the LuKII cell line

(lymphokine/antitumor factor/cancer therapy)

BERISH Y. RUBIN\*†, SYLVIA L. ANDERSON\*, SUSAN A. SULLIVAN\*, BARBARA D. WILLIAMSON†, ELIZABETH A. CARSWELL‡, AND LLOYD J. OLD†

\*Department of Lymphokine Biology, New York Blood Center, New York, NY 10021; and †Laboratory of Experimental Cancer Therapy, Memorial Sloan-Kettering Cancer Center, New York, NY 10021

Contributed by Lloyd J. Old, May 23, 1985

**ABSTRACT** A factor with tumor necrosis factor (TNF) activity produced by the LuKII human lymphoblastoid cell line [designated TNF(LuKII)] was purified sequentially by using controlled-pore glass, lentil lectin-Sepharose, and procion red agarose chromatography, yielding TNF with a specific activity of  $1.5 \times 10^7$  units per mg of protein and an isoelectric point of  $\approx 6.7$ . Purified TNF(LuKII) fractionated by NaDodSO<sub>4</sub>/PAGE under reducing as well as nonreducing conditions was found to contain seven protein bands of  $M_r$  80,000, 70,000, 43,000, 25,000, 23,000, 21,000, and 19,000. The proteins of  $M_r$  80,000 and 70,000 could not be dissociated into lower molecular weight components. Peptide mapping analysis and immunoblotting analysis revealed that the seven protein bands in the purified TNF(LuKII) preparations are related. After fractionation of TNF(LuKII) by NaDodSO<sub>4</sub>/PAGE under reducing conditions, TNF activity was recovered from the regions of  $M_r$  70,000 and 19,000-25,000. Purified human TNF(LuKII) (i) produces hemorrhagic necrosis of Meth A mouse sarcoma in the standard *in vivo* mouse TNF assay; (ii) has the same pattern of reactivity as mouse TNF (cytotoxic/cytostatic/no effect) on a panel of human cancer cell lines; and (iii) has its anticellular effect potentiated by interferon, also a feature of mouse TNF.

The presence of a tumor inhibitory factor in the sera of mice infected with bacillus Calmette-Guérin (BCG) and subsequently injected with endotoxin was reported by Carswell *et al.* (1). Sera from these mice cause necrosis and regression of certain tumors in mice and have a cytotoxic effect on tumor cells *in vitro* (1-5). By using similar methods, a factor with the same *in vivo* and *in vitro* properties can be induced in rats (1) and rabbits (1, 6, 7). The antitumor factor present in the sera of animals sensitized to BCG or other immunopotentiating agents, such as *Corynebacterium parvum* or Zymosan, and then challenged with endotoxin has been termed tumor necrosis factor (TNF). Biochemical studies have indicated that serum TNF activity is associated with both high molecular weight components (4, 8) and components in the  $M_r$  range of 40,000 to 70,000 (3-5, 9, 10).

We have recently reported that human cell lines of hematopoietic origin have the capacity to produce a factor with TNF activity (11). The product of one of the lines (LuKII) was chosen for detailed studies and, according to the following criteria, TNF(LuKII) and mouse TNF have identical properties: (i) mouse L cells made resistant to mouse TNF are resistant to TNF(LuKII), and L cells made resistant to TNF(LuKII) are resistant to mouse TNF; (ii) the anticellular response of a panel of human cell lines to TNF(LuKII) or mouse TNF is indistinguishable and can be potentiated in a synergistic fashion by interferon; and (iii) TNF(LuKII)

causes hemorrhagic necrosis of Meth A sarcomas in the standard *in vivo* TNF assay (11).

In the present study, we report a sequential chromatographic procedure for the purification of TNF(LuKII) and describe biochemical, serological, and biological characteristics of purified TNF(LuKII).

### MATERIALS AND METHODS

**Production of TNF(LuKII).** LuKII cells ( $8 \times 10^5$  cells per ml) in RPMI 1640 medium containing 8% fetal calf serum were incubated with 10 ng of mezerein per ml (L. C. Services, Woburn, MA) for 48 hr. The cells were separated by centrifugation, resuspended in fresh RPMI medium lacking any protein supplement, and incubated for an additional 48 hr. Cells were removed by centrifugation, and the culture media were used as the source of TNF(LuKII).

**In Vitro TNF Assay.** TNF assays were performed in 96-well microtiter plates. Serially diluted fractions were sterilized by ultraviolet radiation and TNF-sensitive L cells were added to each well at a density of  $2 \times 10^4$  cells per well in 100  $\mu$ l. After incubation for 2 days at 37°C, the plates were examined microscopically and the percentage of dead cells was determined. The unitage of the sample was calculated as the reciprocal of the highest dilution that killed 50% of the cells. All TNF assays were run in parallel with a laboratory standard and titers are expressed in laboratory units.

**In Vivo TNF Assay.** The standard Meth A sarcoma assay was performed as described (11).

**Monoclonal Antibody to Human TNF.** BALB/c mice were injected with 1600 units of TNF(LuKII), with a specific activity of  $1.5 \times 10^7$  units/mg. For the initial injection, TNF(LuKII) was mixed with Freund's complete adjuvant (1:1) and was injected subcutaneously. Subsequent injections were given intraperitoneally in the absence of adjuvant. Serum antibody to TNF(LuKII) was determined by an ELISA with TNF(LuKII) bound to polystyrene plates. After nine immunizations over a period of 7 months, the spleen of one mouse with a high titer of TNF(LuKII) antibody was removed and fused with cells of the P<sub>3</sub>U<sub>1</sub> mouse plasmacytoma cell line. Resulting clones were screened for their ability to bind TNF(LuKII) in ELISAs. A hybridoma (designated T1-18) producing antibody reactive with TNF(LuKII) was isolated and subcloned. Media from T1-18 hybridoma cultures served as a source of TNF(LuKII) antibody.

**Affinity Chromatography.** Affinity chromatography procedures were carried out at room temperature and column fractions were collected into polypropylene tubes or bottles. The column matrices used were controlled-pore glass 350 (Electro-Nucleonics, Fairfield, NJ), lentil lectin-Sepharose

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: TNF, tumor necrosis factor.

†To whom reprint requests should be addressed.



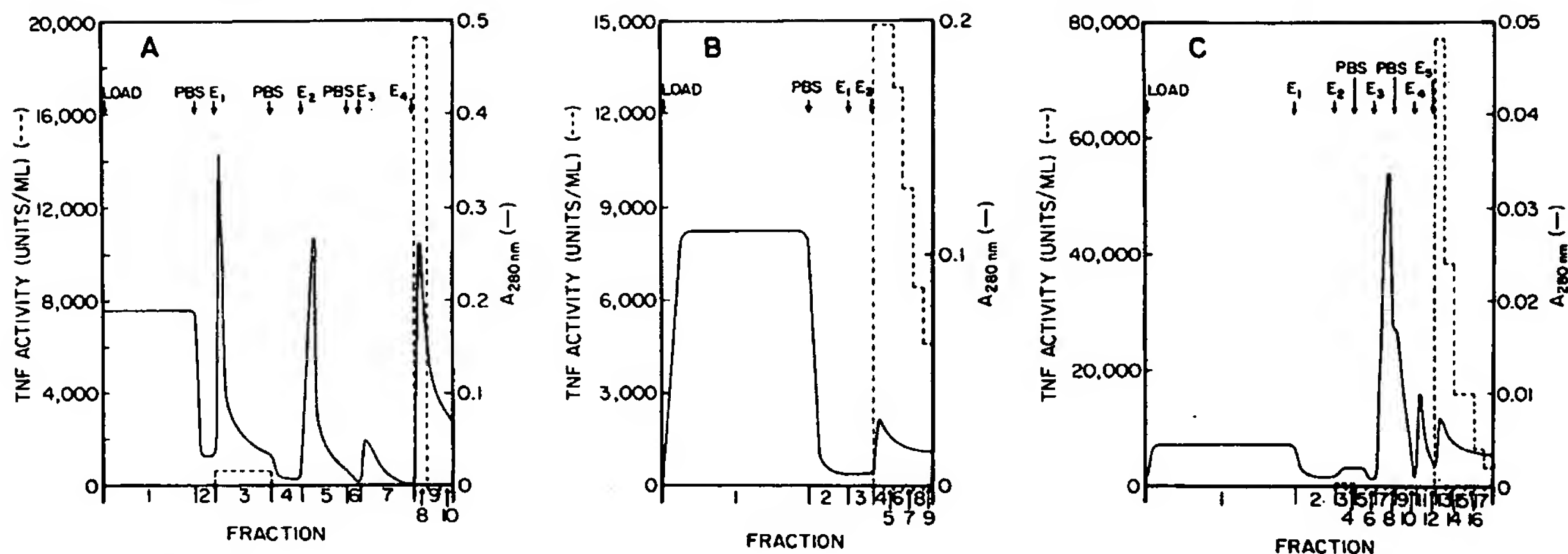


FIG. 1. (A) Controlled-pore glass column chromatography. LuKII culture medium (8 liters) containing 200 units of TNF activity per ml was applied to a controlled-pore glass column (50 ml) equilibrated with phosphate-buffered saline (20 mM sodium phosphate, pH 7.0/0.15 M NaCl) (PBS). The column was washed with the following buffers in sequence: PBS (75 ml); PBS containing 20% ethylene glycol (vol/vol) ( $E_1$ ) (225 ml); PBS (120 ml)/20 mM sodium phosphate, pH 7.0/1.15 M NaCl ( $E_2$ ) (175 ml); PBS (50 ml)/5 mM sodium phosphate, pH 6.8 ( $E_3$ ) (225 ml); and 5 mM sodium phosphate, pH 6.8/5% triethylamine (vol/vol) ( $E_4$ ) (150 ml). Eluted fractions were collected in polypropylene bottles. The material eluted with the  $E_4$  buffer was collected in 50-ml aliquots. (B) Lentil lectin-Sepharose column chromatography. Partially purified TNF(LuKII) (150 ml) eluted from the controlled-pore glass column was loaded onto a lentil lectin-Sepharose column (10 ml) equilibrated with PBS. The column was washed sequentially with PBS (40 ml), PBS/1 M NaCl ( $E_1$ ) (24 ml), and PBS/1 M NaCl/0.2 M methyl- $\alpha$ -D-mannoside ( $E_2$ ) (60 ml). The material eluted with the methyl- $\alpha$ -D-mannoside-containing buffer was collected in 10-ml aliquots. (C) Procion red agarose column chromatography. Partially purified TNF(LuKII) (60 ml) eluted from the lentil lectin-column was diluted 1:1 with PBS and loaded onto a procion red agarose column (4 ml) equilibrated with 20 mM sodium phosphate, pH 6.8/0.65 M NaCl (PBS/0.5 M NaCl). The column was washed with the following buffers in sequence: PBS/0.5 M NaCl ( $E_1$ ) (30 ml); PBS/1 M NaCl ( $E_2$ ) (8 ml); PBS (8 ml); PBS/50% ethylene glycol (vol/vol) ( $E_3$ ) (8 ml); PBS (8 ml); 0.1 M Tris-HCl, pH 9.4/0.1 M NaCl ( $E_4$ ) (8 ml); and 0.1 M Tris-HCl, pH 9.4/0.1 M arginine ( $E_5$ ) (24 ml). The material eluted with the 0.1 M Tris-HCl, pH 9.4/0.1 M arginine buffer was collected in 4-ml aliquots.

(Pharmacia), and procion red agarose (Bethesda Research Laboratories).

**Protein Determinations.** Protein determinations were carried out with the Bio-Rad dye reagent (Bio-Rad) using bovine serum albumin as a standard.

**Radiolabeling of TNF(LuKII).** TNF(LuKII) was labeled with  $^{125}\text{I}$  using 1,3,4,6-tetrachloro-3 $\alpha$ ,6 $\alpha$ -diphenylglycouril (Iodo-Gen, Pierce). Polypropylene tubes were coated with 100  $\mu\text{g}$  of Iodo-Gen (dissolved in chloroform) by evaporation of the solvent. A 2-ml sample of TNF(LuKII) (50,000 units/ml) with a specific activity of  $1.5 \times 10^7$  units per mg of protein was incubated for 25 min at room temperature in an Iodo-Gen-coated tube containing 2 mCi of  $^{125}\text{I}$  (1 Ci = 37 GBq). The labeled protein was then separated from the unbound  $^{125}\text{I}$  by using a P-4 column (Bio-Rad) equilibrated with phosphate-buffered saline containing 50  $\mu\text{g}$  of cytochrome *c* per ml. The iodinated material eluted in the void volume of the column was divided into aliquots and stored at  $-80^\circ\text{C}$ .

**NaDodSO<sub>4</sub>/PAGE.** NaDodSO<sub>4</sub>/PAGE was carried out in 18-cm slab gels according to published methods (12).

**Isoelectrofocusing.** Isoelectrofocusing was performed by using Ampholine Pagplates (pH 3.5–9.5) (LKB). The gels were run at 30 W for 1.5 hr, at which time the pH gradient was measured and the gel was sliced into 18 equal pieces. The gel fractions were incubated for 18 hr in Eagle's minimum

essential medium (ME medium) containing 10% fetal calf serum and fractions were assayed for TNF *in vitro*.

**Peptide Mapping Analysis.** A  $^{125}\text{I}$ -labeled preparation of purified TNF(LuKII) was fractionated by NaDodSO<sub>4</sub>/PAGE and individual bands localized by autoradiography were cut from the gel and treated with L-1-tosylamido-2-phenylethyl chloromethyl ketone-treated trypsin or *N*- $\alpha$ -tosyllysine chloromethyl ketone-treated chymotrypsin. Digested fractions were analyzed according to the methods of Elder *et al.* (13).

**Immunoblotting Analysis.** Immunoblotting was performed essentially as described (14). Briefly, preparations of purified TNF(LuKII) were transferred to nitrocellulose paper overnight at 100 mA. After incubation of the nitrocellulose paper in buffer containing bovine serum albumin, the paper was exposed for 2 hr to 40 ml of T1-18 antibody-containing culture medium. The nitrocellulose paper was then washed extensively and incubated overnight in 10 mM Tris-HCl, pH 7.4/0.9% NaCl/5% bovine serum albumin/ $^{125}\text{I}$ -labeled rabbit anti-mouse IgG. The nitrocellulose paper was further washed and exposed to x-ray film.

## RESULTS

**Purification of TNF(LuKII).** Controlled-pore glass beads bound all TNF activity from LuKII culture fluids. After washing with several buffers in sequence, TNF activity was

Table 1. Purification of TNF(LuKII)

Column	Load		Recovery		
	Units	Specific activity, units/mg	Units	Specific activity, units/mg	% recovery
Controlled-pore glass	$1.6 \times 10^6$	$5.3 \times 10^3$	$9.6 \times 10^5$	$3.8 \times 10^5$	60
Lentil lectin-Sepharose	$9.6 \times 10^5$	$3.8 \times 10^5$	$6.3 \times 10^5$	$1.3 \times 10^6$	39
Procion red agarose	$6.3 \times 10^5$	$1 \times 10^6$	$6.3 \times 10^5$	$1.5 \times 10^7$	39
					-Fold purification
					72
					245
					2830



eluted with a 5 mM sodium phosphate buffer (pH 6.8) containing 5% triethylamine (Fig. 1A). The eluted TNF was then applied to a lentil lectin-Sepharose column, which was washed first with phosphate-buffered saline and then with 0.02 M sodium phosphate buffer (pH 6.8) containing 1.15 M NaCl (buffer A). TNF activity was eluted from this column with buffer A containing 0.2 M methyl- $\alpha$ -D-mannoside (Fig. 1B). All TNF activity bound to the lentil lectin-Sepharose column and 39% of the activity was recovered in the methyl- $\alpha$ -D-mannoside-containing buffer. (Further washing of the column with buffer containing 50% ethylene glycol elutes only a small amount of additional TNF activity.) TNF from the lentil lectin column was then diluted 1:1 with phosphate-buffered saline and loaded onto a procion red agarose column. The column was washed sequentially with several buffers that remove protein having no TNF activity. The column was then washed with 0.1 M Tris-HCl, pH 9.4/0.1 M arginine. TNF activity was eluted with this buffer, yielding TNF with a specific activity of  $1.5 \times 10^7$  units per mg of protein. Table 1 summarizes the purification scheme for TNF(LuKII) with specific activities of the resulting fractions.

**Biochemical Characterization of Purified TNF(LuKII).** Isoelectric focusing of purified TNF(LuKII) indicates an isoelectric point of  $\approx 6.7$  (Fig. 2).  $^{125}$ I-labeled TNF(LuKII) with a specific activity of  $1.5 \times 10^7$  units per mg of protein was analyzed by NaDodSO<sub>4</sub>/PAGE and found to contain seven protein bands with  $M_r$  values of 80,000, 70,000, 43,000, 25,000, 23,000, 21,000, and 19,000 (Fig. 3). The same seven protein bands were observed when nonlabeled purified TNF(LuKII) was fractionated by NaDodSO<sub>4</sub>/PAGE and examined by silver staining. The proteins of  $M_r$  80,000 and 70,000 were eluted from the gels and reanalyzed by NaDodSO<sub>4</sub>/PAGE. They migrated once again to the  $M_r$  70,000–80,000 region, and no lower molecular weight components were observed. In further experiments, purified TNF(LuKII) was boiled in NaDodSO<sub>4</sub>, urea, and 2-mercaptoethanol, and the same characteristic seven bands were found.

To determine which bands in TNF(LuKII) showed TNF activity, parallel samples of purified TNF(LuKII), one  $^{125}$ I-labeled and one unlabeled, were treated with 0.1% NaDodSO<sub>4</sub>/0.1 M 2-mercaptoethanol and fractionated by NaDodSO<sub>4</sub>/PAGE. After electrophoresis, the lane containing the unlabeled material was cut into 4.4-mm slices and the proteins were eluted from each slice by overnight incubation at 4°C in ME medium containing fetal calf serum. The parallel lane containing  $^{125}$ I-labeled TNF(LuKII) was dried immediately after electrophoresis and protein bands were located by autoradiography. As seen in Fig. 4, TNF activity was recovered from the gel at  $M_r$  values of 70,000 and 19,000–25,000, corresponding to  $^{125}$ I-labeled protein bands at these posi-

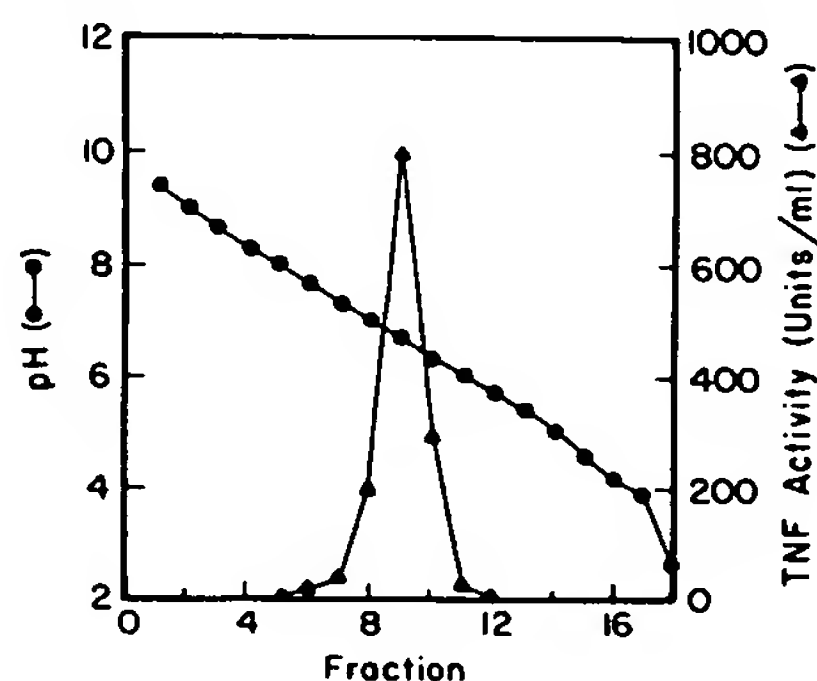


FIG. 2. Isoelectrofocusing of TNF(LuKII). A 60- $\mu$ l sample of purified TNF(LuKII) containing 1500 units was applied to an ampholine gel (pH 3.5–9.5).

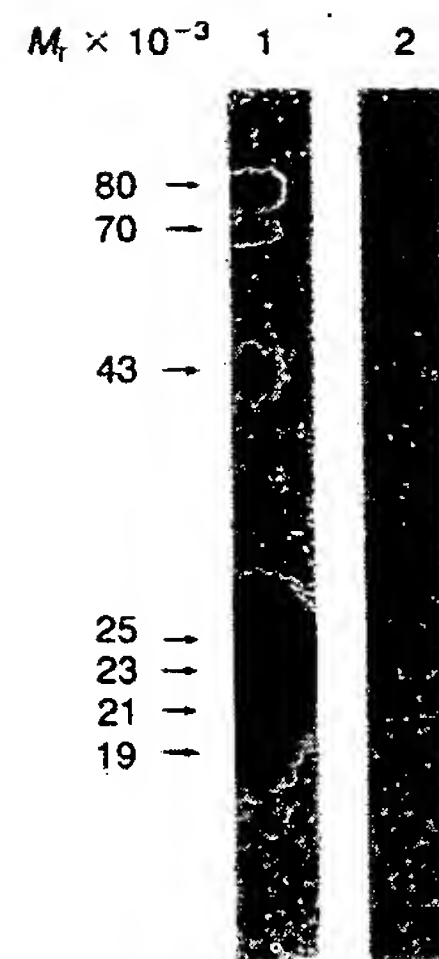


FIG. 3. NaDodSO<sub>4</sub>/PAGE of purified  $^{125}$ I-labeled TNF(LuKII). TNF(LuKII) was iodinated and fractionated by NaDodSO<sub>4</sub>/PAGE. Autoradiographs were developed for 18 hr (lane 1) and 0.5 hr (lane 2). The following proteins provided  $M_r$  markers: myosin ( $M_r$ , 200,000),  $\beta$ -galactosidase ( $M_r$ , 130,000), phosphor-ylase b ( $M_r$ , 94,000), bovine serum albumin ( $M_r$ , 67,000), ovalbumin ( $M_r$ , 43,000),  $\alpha$ -chymotrypsinogen ( $M_r$ , 25,700),  $\beta$ -lactoglobulin ( $M_r$ , 18,400), lysozyme ( $M_r$ , 14,300), and cytochrome c ( $M_r$ , 12,300).

tions. TNF(LuKII) samples that were not exposed to 2-mercaptoethanol before NaDodSO<sub>4</sub>/PAGE also showed TNF activity at  $M_r$  values of 70,000 and 19,000–25,000.

To examine the relationships among the various protein bands in purified TNF(LuKII) preparations, two-dimensional chymotryptic and tryptic peptide mapping analyses were performed. As seen in Fig. 5a, the chymotryptic peptide maps demonstrate that the proteins of  $M_r$  43,000, 25,000, 23,000, 21,000, and 19,000 are related and the proteins of  $M_r$  80,000 and 70,000 are related. To examine the relationship of the larger molecular weight proteins to the smaller proteins, chymotryptic digests of the  $M_r$  70,000 and 25,000 proteins containing equal amounts of radioactivity were mixed and analyzed. As seen in Fig. 5b, three of the fragments (termed A, B, and C) generated by digestion of the  $M_r$  25,000 protein migrate to the same position as three fragments generated by digestion of the  $M_r$  70,000 protein. A similar analysis was carried out using trypsin as the proteolytic enzyme. The results also indicate that the seven distinct forms are closely related.

Further evidence for the relationships among the various proteins in purified TNF(LuKII) comes from immunoblotting analysis with T1-18 monoclonal antibody to TNF(LuKII). Fig. 6 shows that the antibody reacts with the  $M_r$  43,000 and the  $M_r$  19,000–25,000 components.

#### Biological Characteristics of Purified TNF(LuKII). Limulus

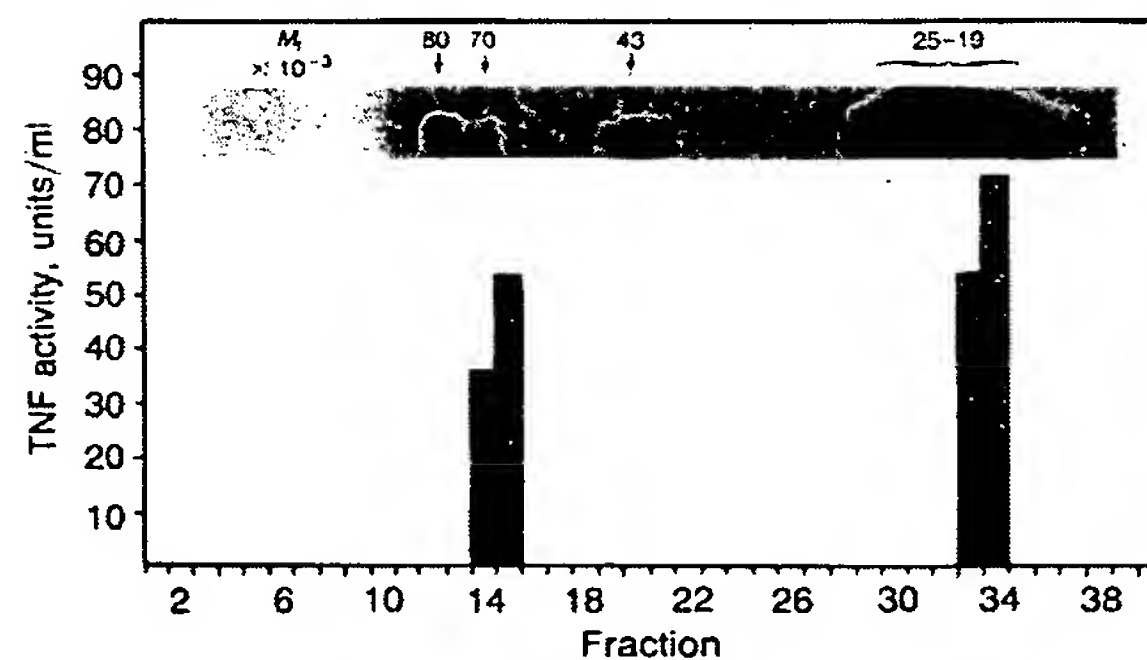


FIG. 4. Recovery of TNF activity after NaDodSO<sub>4</sub>/PAGE fractionation of TNF(LuKII). A sample of TNF(LuKII) containing 6000 units adjusted to contain 0.1% NaDodSO<sub>4</sub> and 0.1 M 2-mercaptoethanol was applied to a 12% polyacrylamide gel. After electrophoresis, the gel was sliced and activity was eluted and assayed. In an adjacent lane,  $^{125}$ I-labeled TNF(LuKII) was fractionated and autoradiographed to determine the molecular weight of the TNF(LuKII) active fractions.

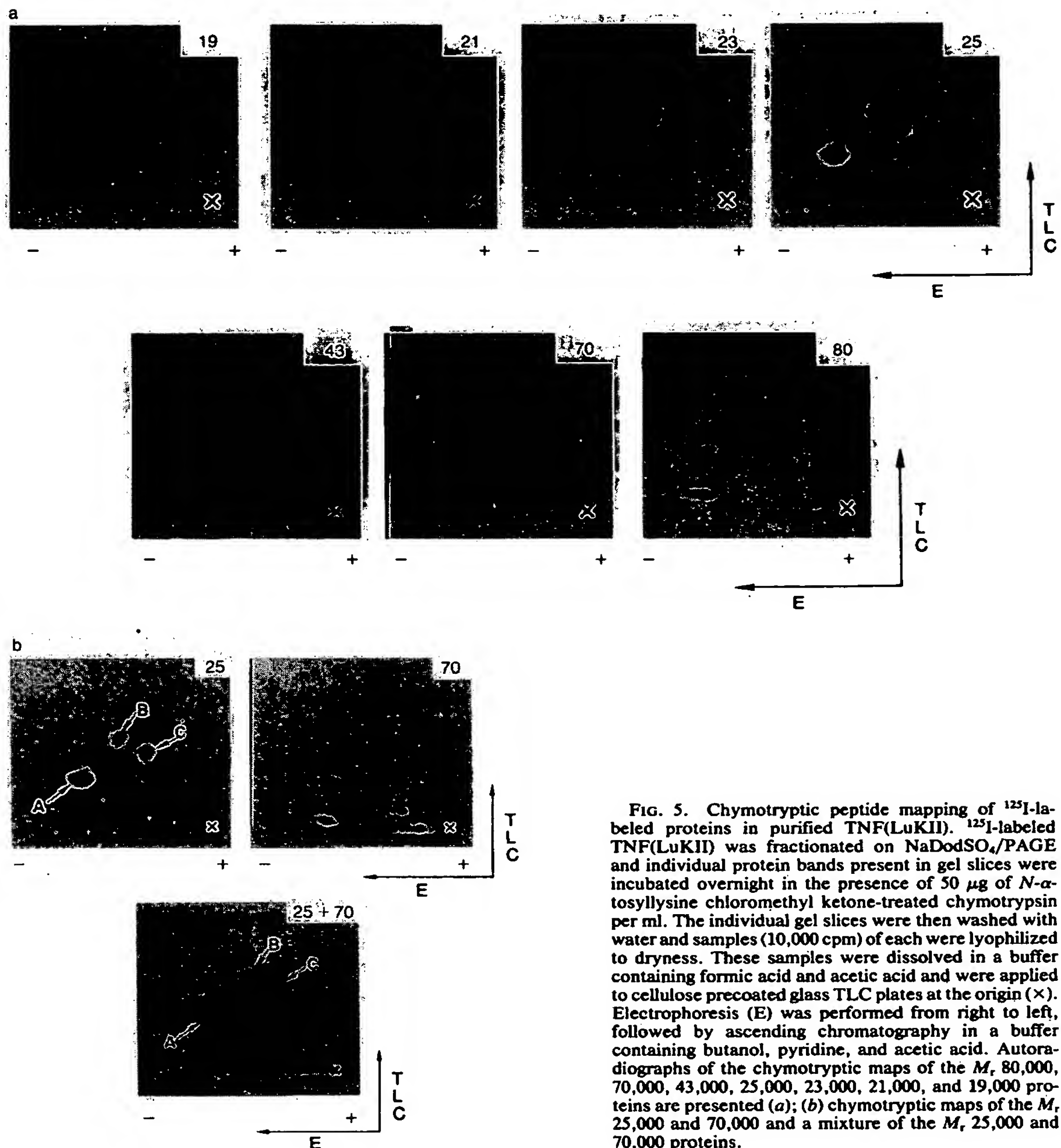


FIG. 5. Chymotryptic peptide mapping of  $^{125}\text{I}$ -labeled proteins in purified TNF(LuKII).  $^{125}\text{I}$ -labeled TNF(LuKII) was fractionated on NaDodSO<sub>4</sub>/PAGE and individual protein bands present in gel slices were incubated overnight in the presence of 50  $\mu\text{g}$  of *N*- $\alpha$ -tosyllysine chloromethyl ketone-treated chymotrypsin per ml. The individual gel slices were then washed with water and samples (10,000 cpm) of each were lyophilized to dryness. These samples were dissolved in a buffer containing formic acid and acetic acid and were applied to cellulose precoated glass TLC plates at the origin (X). Electrophoresis (E) was performed from right to left, followed by ascending chromatography in a buffer containing butanol, pyridine, and acetic acid. Autoradiographs of the chymotryptic maps of the  $M_r$  80,000, 70,000, 43,000, 25,000, 23,000, 21,000, and 19,000 proteins are presented (a); (b) chymotryptic maps of the  $M_r$  25,000 and 70,000 and a mixture of the  $M_r$  25,000 and 70,000 proteins.

tests of purified TNF(LuKII) indicate 25 ng of endotoxin per ml. TNF(LuKII) causes hemorrhagic necrosis of Meth A sarcoma after intratumoral or intravenous injection and total tumor regression has been observed in some treated mice. L-cell lines made resistant to mouse TNF or to partially purified TNF(LuKII) are resistant to purified TNF(LuKII). With the panel of human cell lines studied by Williamson and coworkers (11), purified TNF(LuKII) showed the same pattern of reactivity (cytotoxic/cytostatic/no effect) as mouse TNF and partially purified TNF(LuKII). In addition, purified TNF(LuKII) and interferon showed synergistic cytotoxic activity for human tumor cells, similar to what has

previously been reported for partially purified TNF(LuKII) and mouse TNF (12).

#### DISCUSSION

We have recently described the production, characterization, and biological properties of a human factor with TNF activity from the LuKII lymphoblastoid cell line (11). The further purification and characterization of this factor, designated TNF(LuKII), is the subject of this report. A protocol for the purification of TNF(LuKII) has been developed that yields both good recoveries of TNF and material with high specific activity. This purification protocol allows active fractions

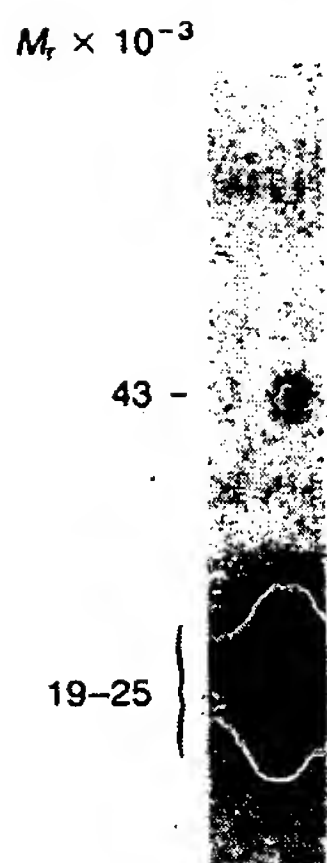


FIG. 6. Immunoblotting analysis of TNF(LuKII) with T1-18 mouse monoclonal antibody. A sample of TNF(LuKII) containing 10,000 units was fractionated by NaDodSO<sub>4</sub>/PAGE. Fractionated proteins were transferred to a nitrocellulose membrane and processed as described.

eluted from one column to be applied either directly or after dilution onto the next column, thereby eliminating any need for dialysis and thus avoiding the losses associated with dialysis. TNF(LuKII) has a  $M_r$  of 70,000 by gel filtration under nonreducing conditions and an isoelectric point of 6.7. Examination of purified TNF(LuKII) by NaDodSO<sub>4</sub>/PAGE under reducing as well as nonreducing conditions revealed the presence of seven protein bands ranging from  $M_r$  19,000 to 80,000. The  $M_r$  80,000 and 70,000 proteins could not be dissociated into smaller molecular weight components even after boiling in NaDodSO<sub>4</sub>/2-mercaptoethanol/urea. Fractions from NaDodSO<sub>4</sub>/PAGE were assayed for TNF activity and proteins in the  $M_r$  70,000 and 19,000–25,000 region had TNF activity. Using peptide mapping analysis, we found that the seven proteins present in purified TNF(LuKII) were related. Immunoblotting analysis with monoclonal antibody to TNF(LuKII) showed shared determinants on the  $M_r$  43,000 and 19,000–25,000 proteins. Antibody did not react with the higher molecular weight forms, even though these have been shown to be related to the  $M_r$  43,000 and lower molecular weight components. This could be due to the inaccessibility of the determinant on the  $M_r$  70,000 and 80,000 species. Thus, our analysis indicates that there are a number of structurally related proteins in purified TNF(LuKII) and that TNF activity is associated with nondissociable high molecular weight and low molecular weight forms. We conclude that the seven proteins in our purified TNF(LuKII) are either the products of related genes or products of a single gene that undergoes extensive processing.

TNF(LuKII) has the full range of biological activities associated with mouse TNF. It produces hemorrhagic necrosis of Meth A sarcoma in the standard TNF assay and cannot be distinguished from mouse TNF in its pattern of reactivity on a large panel of human cancer cell lines. In addition, L cells made resistant to mouse TNF are resistant to TNF(LuKII), and cells made resistant to TNF(LuKII) are resistant to mouse TNF. Recent work has indicated that there are surface receptors for TNF on TNF-sensitive cells (unpublished data). Competitive binding studies showed that mouse TNF and TNF(LuKII) compete for the same receptor.

The relationship between TNF(LuKII) and lymphotoxin (15, 16) is unclear. While they share certain properties, such as their ability to kill mouse L cells, their affinity for lentil lectin, their multiple forms, and, in some cases, their cellular origin, they differ in certain biochemical properties. Aggarwal and coworkers (17, 18) have observed that lympho-

toxin from the RPMI 1788 lymphoblastoid cell line has a  $M_r$  value of 25,000 and 20,000 under reducing conditions, whereas TNF(LuKII) exists in several molecular weight forms, some of which do not dissociate under reducing conditions. Granger *et al.* (15) have also reported that species of lymphotoxin exist in higher molecular weight forms. It seems likely that there is a family of cytotoxic factors with TNF activity and that lymphotoxin and TNF(LuKII) are members of this family. This would be comparable to the interferon system in which there are three major species, all having antiviral activities, but each being coded for by a different gene with varying degrees of homology. As in the interferon field, the cloning of factors with TNF activity will provide the basis for distinguishing various molecules of this family. The recent cloning of human TNF (19–21) and human lymphotoxin (22) has clarified the relationship between these two molecules. Recombinant TNF and lymphotoxin are distinct molecules of  $M_r$  17,000 and 18,600 that share considerable sequence homology. Both kill L cells and have tumor necrosis activity. In light of the observation that TNF(LuKII) exists in higher molecular weight forms, it seems unlikely that all forms of TNF have been cloned.

This research was supported in part by grants from the National Cancer Institute (CA-38661 and CA-08748).

1. Carswell, E. A., Old, L. J., Kassel, R. L., Green, S., Fiore, N. & Williamson, B. (1975) *Proc. Natl. Acad. Sci. USA* 72, 3666–3670.
2. Helson, L., Green, S., Carswell, E. A. & Old, L. J. (1975) *Nature (London)* 258, 731–732.
3. Mannel, D. N., Meltzer, M. S. & Mergenhagen, S. E. (1980) *Infect. Immun.* 28, 204–211.
4. Kull, F. C. & Cuatrecasas, P. (1981) *J. Immunol.* 126, 1279–1283.
5. Haranaka, K. & Satomi, N. (1981) *Jpn. J. Exp. Med.* 51, 191–194.
6. Matthews, N. & Watkins, J. F. (1978) *Br. J. Cancer* 38, 302–309.
7. Ostrove, J. M. & Gifford, G. E. (1979) *Proc. Soc. Exp. Biol. Med.* 160, 354–358.
8. Green, S., Dobrjansky, A., Carswell, E. A., Kassel, R. L., Old, L. J., Fiore, N. & Schwartz, M. K. (1976) *Proc. Natl. Acad. Sci. USA* 73, 381–385.
9. Matthews, N., Ryley, H. C. & Neale, M. L. (1980) *Br. J. Cancer* 42, 416–422.
10. Ruff, M. R. & Gifford, G. E. (1980) *J. Immunol.* 125, 1671–1677.
11. Williamson, B. D., Carswell, E. A., Rubin, B. Y., Prendergast, J. S. & Old, L. J. (1983) *Proc. Natl. Acad. Sci. USA* 80, 5397–5401.
12. Laemmli, U. K. (1970) *Nature (London)* 227, 680–685.
13. Elder, J. H., Jensen, F. C., Bryant, M. L. & Lerner, R. A. (1977) *Nature (London)* 267, 23–28.
14. Burnette, W. N. (1981) *Anal. Biochem.* 112, 195–203.
15. Granger, G. A., Yamamoto, R. S., Fair, D. S. & Hiserodt, J. C. (1978) *Cell. Immunol.* 38, 388–402.
16. Rosenau, W. (1981) *Int. J. Immunopharmacol.* 3, 1–8.
17. Aggarwal, B. B., Moffat, B. & Harkins, R. N. (1984) *J. Biol. Chem.* 259, 686–691.
18. Aggarwal, B. B., Henzel, W. J., Moffat, B., Kohr, W. J. & Harkins, R. N. (1985) *J. Biol. Chem.* 260, 2334–2344.
19. Pennica, D., Nedwin, G. E., Hayflick, J. S., Seeburg, P. H., Derynck, R., Palladino, M. A., Kohr, W. J., Aggarwal, B. B. & Goeddel, D. V. (1984) *Nature (London)* 312, 724–729.
20. Shirai, T., Yamaguchi, H., Ito, H., Todd, C. W. & Wallace, R. B. (1985) *Nature (London)* 313, 803–806.
21. Wang, A. M., Creasey, A. A., Ladner, M. B., Lin, L. S., Strickler, J., Van Arsdell, J. N., Yamamoto, R. & Mark, D. F. (1985) *Science* 228, 149–154.
22. Gray, P. W., Aggarwal, B. B., Benton, C. V., Bringman, T. S., Henzel, W. J., Jarrett, J. A., Leung, D. W., Moffat, B., Ng, P., Svedersky, L. P., Palladino, M. A. & Nedwin, G. E. (1984) *Nature (London)* 312, 721–724.



## Human tumor necrosis factor produced by human B-cell lines: Synergistic cytotoxic interaction with human interferon

(lymphokine/antitumor factor/cancer therapy)

BARBARA D. WILLIAMSON, ELIZABETH A. CARSWELL, BERESH Y. RUBIN, JAY S. PRENDERGAST,  
AND LLOYD J. OLD\*

Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021

Contributed by Lloyd J. Old, May 31, 1983

**ABSTRACT** Human cell lines of hematopoietic origin were tested for production of tumor necrosis factor (TNF). B-cell lines transformed by Epstein-Barr virus release a factor (referred to as hTNF) that is cytotoxic for mouse L cells sensitive to mouse TNF but not for L cells resistant to mouse TNF. Exposure to 4 $\beta$ -phorbol 12 $\beta$ -myristate 13 $\alpha$ -acetate augmented production of hTNF. hTNF activity was not found in supernatants of cell lines of T-cell, monocytic, or promyelocytic origin. Partially purified hTNF has a molecular weight of approximately 70,000, has no interferon activity, is acid labile, is destroyed by heating at 70°C for 1 hr, induces cross-resistance to mouse TNF *in vitro*, and causes hemorrhagic necrosis of Meth A mouse sarcoma in the standard *in vivo* mouse TNF assay. Tests with a panel of 23 human cancer cell lines showed that hTNF is cytotoxic for 7 cell lines, cytostatic for 5, and has no effect on 11. Comparative studies with human  $\alpha$ ,  $\beta$ , and  $\gamma$  interferons indicated that sensitivity to hTNF and interferon can be distinguished. Combined treatment with hTNF and  $\alpha$  or  $\gamma$  interferon resulted in a synergistic cytotoxic effect.

Tumor necrosis factor (TNF) was recognized by Carswell *et al.* (1) during a study of the antitumor activity of serum from mice infected with bacillus Calmette-Guérin (BCG) and subsequently injected with endotoxin. Serum from these mice contains a factor that induces hemorrhagic necrosis of certain mouse sarcomas *in vivo* and has cytotoxic/cytostatic effects on mouse and human tumor cells *in vitro* (1-5). Serum from mice singly treated with BCG or endotoxin did not have these properties. Other agents, such as *Corynebacterium parvum* or zymosan, that cause hyperplasia of the reticuloendothelial system and increase sensitivity to endotoxin lethality, can substitute for BCG in priming for TNF release. Endotoxin, however, appears to be unique in its ability to elicit TNF release (1, 4-6). By using these methods for TNF production, factors with TNF-like activities have also been found in rats (1) and rabbits (1, 7, 8).

Biochemical characterization has shown that mouse serum TNF exists in at least two forms: a 150,000  $M_r$  form (4, 9) and a 40,000-60,000  $M_r$  form (3-5). TNF in rabbit serum has been reported to have a molecular weight of 39,000 (10) and 67,000 (11). Our studies have indicated that both *in vivo* and *in vitro* activities of mouse TNF appear to be a property of the same molecule. The cellular source of TNF in the mouse was initially assumed to be the macrophage, because the agents used to prime for TNF production cause massive hyperplasia of macrophages in liver and spleen (1). From studies of macrophage-rich cell populations *in vitro*, Matthews (12) and Mannel *et al.* (13) reached a similar conclusion with regard to the source of mouse and rabbit TNF. Direct evidence that macrophages are at least one cell type in the mouse capable of producing TNF comes from stud-

ies with cloned lines of mouse histiocytes (ref. 13; unpublished data). These cells constitutively produce low levels of TNF that are greatly increased after exposure to endotoxin.

In the present study we have examined the capacity of human cell lines of hematopoietic origin to produce a TNF-like factor.

### MATERIALS AND METHODS

**Cell Lines.** See Tables 1 and 2. LuKII cells were obtained from W. Stewart (14). B-cell lines transformed by Epstein-Barr virus were derived from patients with melanoma (15). These and other cell lines were from our cell bank or from the collection of Jørgen E. Fogh or Peter Ralph of Sloan-Kettering Institute.

**In Vitro TNF Assay.** TNF-sensitive and -resistant L-M cells were derived from clone 929 mouse L cells (American Type Culture Collection). TNF activity was assayed by adding equal volumes of serially diluted samples to wells (24-well Costar plates) seeded 2-3 hr previously with  $5 \times 10^4$  trypsinized L cells and determining cell death at 48 hr by phase-contrast microscopy (1).

**In Vivo TNF Assay.** (BALB/c  $\times$  C57BL/6) $F_1$  female mice were injected intradermally with  $5 \times 10^5$  Meth A BALB/c sarcoma cells. After 7 days (tumor size approximately 7 mm average diameter), mice received a single intravenous injection of the TNF preparation. After 24 hr, tumor hemorrhagic necrosis was scored according to ref. 1.

**Interferon (IFN) Assays.** Human IFN activity was measured by inhibition of the cytopathic effect of vesicular stomatitis virus on WISH or GM2767 cells (16) and compared against an international standard in the case of IFN- $\alpha$  and the laboratory standard in the case of IFN- $\gamma$ .

**Mouse TNF.** A standard lot of partially purified mouse serum TNF with a specific activity of  $2 \times 10^4$  units of TNF per mg of protein was used in these studies. (A unit of TNF is defined as the amount of protein causing killing of 50% of the L cells in the standard *in vitro* TNF assay.) This preparation of mouse TNF lacked IFN activity.

**Source of IFNs.** Human IFNs were obtained from the following sources: Recombinant IFN- $\alpha$  (Hoffmann-La Roche),  $2-4 \times 10^6$  units/mg of protein; leukocyte-derived IFN- $\alpha$  (prepared for the Sloan-Kettering Institute by Kocher Laboratory, Berne, Switzerland),  $0.64 \times 10^6$  units/mg of protein; leukocyte-derived IFN- $\gamma$ ,  $1 \times 10^7$  units/mg of protein; fibroblast-derived IFN- $\beta$  (Roswell Park Memorial Inst.),  $1 \times 10^7$  units/mg of protein.

Abbreviations: TNF, tumor necrosis factor; hTNF, human TNF; BCG, bacillus Calmette-Guérin; PMA, 4 $\beta$ -phorbol 12 $\beta$ -myristate 13 $\alpha$ -acetate; IFN, interferon.

\*To whom reprint requests should be addressed.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.



**Screening of Human Cells of Hematopoietic Origin for TNF Production.** Cells were cultured at  $5 \times 10^5$  per ml in RPMI 1640 medium containing 8% fetal calf serum with or without  $4\beta$ -phorbol  $12\beta$ -myristate  $13\alpha$ -acetate (PMA) (Sigma) at 10 ng/ml. After incubation for 48 hr at 37°C, the cells were collected by centrifugation and resuspended at  $1 \times 10^6$  per ml in medium without PMA for an additional 48 hr. Culture supernatants were freed of cells by centrifugation and frozen at -20°C before screening 1/2 diluted supernatants for TNF activity by the L cell assay.

**Preparation and Concentration of hTNF from LuKII Cells.** LuKII cells were cultured at  $5 \times 10^5$  per ml in RPMI 1640 medium containing 5% fetal calf serum and PMA at 10 ng/ml and incubated for 48 hr at 37°C. At 48 hr, cells were collected by centrifugation and resuspended at  $8-10 \times 10^5$  per ml in serum-free R-ITS PREMIX (insulin, transferrin, selenium serum substitute, Collaborative Research, Waltham, MA) and 2 nM ethanolamine and incubated an additional 48 hr. Supernatants were spun free of cells and frozen at -20°C. LuKII supernatants were concentrated by Amicon stirred cells (PM 10 membrane) and applied to a DEAE-Sephadex A-50 column ( $40 \times 2.6$  cm) equilibrated with 0.05 M Tris-HCl/0.15 M NaCl buffer, pH 7.3. Fractions (5 ml) were eluted with the same buffer and hTNF-active fractions were pooled and concentrated by Amicon cells. Molecular weight was estimated by Sephacryl S-200 column chromatography.

**Effect of hTNF and IFN on Human Cell Lines.** Cells were trypsinized, rinsed twice, plated at  $4-5 \times 10^4$  per well (24-well Costar plates) in minimal essential medium containing 8% fetal calf serum and incubated at 37°C. After 16-24 hr, the medium was aspirated and 1-ml dilutions containing DEAE-fractionated hTNF, IFN, or hTNF and IFN in medium were added. Total cell counts (viable and nonviable) were determined at 3, 5, and 7 days by phase-contrast microscopy. Cultures were re-fed on day 5 with fresh medium containing the same concentrations of hTNF and IFN.

## RESULTS

**TNF Production by Human Cell Lines of Hematopoietic Origin.** Supernatants from the cell lines listed in Table 1 were screened for TNF production by using the L cell assay system developed for the detection of mouse TNF (1, 6). Supernatants were tested on TNF-sensitive L cells [designated  $L_{(S)}$ ] and on the TNF-resistant L cell line [designated  $L_{(R)}$ ]. The  $L_{(R)}$  line was developed by repeated passage of sensitive L cells in medium containing mouse TNF. Table 1 shows that cell lines of B-cell origin produce a factor with TNF-like activity—e.g., cytotoxic for  $L_{(S)}$  but not  $L_{(R)}$ —and that PMA is necessary for efficient TNF production. No TNF activity was detected in the supernatants of six T-cell or three monocytic cell lines (with or without PMA stimulation). The LuKII line consistently produced high levels of TNF after PMA induction and was therefore selected as the source for subsequent studies of TNF production and characterization.

**Fractionation and Characterization of Human TNF (hTNF).** Concentrated samples of LuKII supernatants were fractionated on DEAE-Sephadex A-50 columns and individual fractions were assayed on  $L_{(S)}$  and  $L_{(R)}$  cells. TNF activity was detected in the initial flow-through fractions, and peak fractions were pooled and concentrated. The specific activity of these preparations was  $2-5 \times 10^4$  units/mg of protein (a 25-fold increase over the initial supernatants), and these preparations were used for characterization. The molecular weight of hTNF was estimated to be approximately 70,000. hTNF activity was destroyed by a 12-hr exposure to low or high pH: 90% of the activity was lost at

Table 1. Screening of human cell lines of hematopoietic origin for TNF production

Cell line	Cell viability (%) after exposure to culture supernatants			
	$L_{(S)}$ cells		$L_{(R)}$ cells	
	No PMA	PMA	No PMA	PMA
<b>B cell</b>				
BE*	47	20	100	100
BI*	57	43	100	98
CL*	95	37	100	97
CW*	28	8	100	100
DE*	45	18	100	100
DM*	18	11	100	100
DS*	88	38	100	100
EJ*	27	16	100	100
EL*	70	16	100	100
EQ*	15	11	100	100
ER*	47	22	100	100
FC*	19	8	100	100
FD*	44	16	97	100
FG*	80	30	100	97
ARH-77	68	12	100	100
DAUDI	97	97	NT	NT
LuKII	20	13	100	100
RPMI 1788	78	35	100	100
RPMI 8226	96	84	100	86
RPMI 8866	32	10	100	98
SK-LY-16	98	100	100	100
SK-LY-18	99	100	100	98
ARA-10	77	39	100	97
BALL-1	100	100	100	100
<b>T cell</b>				
CCRF-CEM	100	100	100	100
P-12	95	94	99	98
MOLT-4	100	100	99	96
T-45	100	100	100	100
HPB-ALL	100	97	100	100
TALL-1	96	94	96	96
<b>Monocyte</b>				
J-111	100	100	99	98
THP	96	80	99	99
U-937	100	100	99	99
<b>Promyelocyte</b>				
HL-60	96	97	99	98

NT, not tested.

\* Epstein-Barr virus-transformed cell line derived from peripheral blood.

pH 2.0 and 45% of the activity was lost at pH 10. Activity was stable over a pH range of 6 to 8. With regard to heat stability, hTNF was destroyed at 70°C for 60 min but was stable at 56°C for 60 min. Activity was preserved over long periods of storage at -78 or -20°C. IFN assays demonstrated the presence of 100 units of IFN- $\alpha$  per ml in the unfractionated supernatant from PMA-stimulated LuKII cells. No IFN activity was detected in hTNF preparations after DEAE-Sephadex chromatography. To investigate further the relationship of mouse TNF to hTNF, L cells were made resistant to hTNF by repeated passage in hTNF-containing medium. hTNF-resistant cells showed complete cross-resistance to mouse TNF. DEAE-fractionated hTNF was assayed by the standard *in vivo* mouse TNF assay (1) and found to cause necrosis of Meth A tumors. Three hundred micrograms of the hTNF preparations produced +++ necrosis in 1/5 mice and ++ in 4/5 mice; 150  $\mu$ g produced ++ necrosis in 5/7 mice and + in 2/7 mice; 75  $\mu$ g produced ++ necrosis in 1/5 mice and + in 4/5 mice. No necrosis was observed in control mice after injection of comparable DEAE-Sephadex

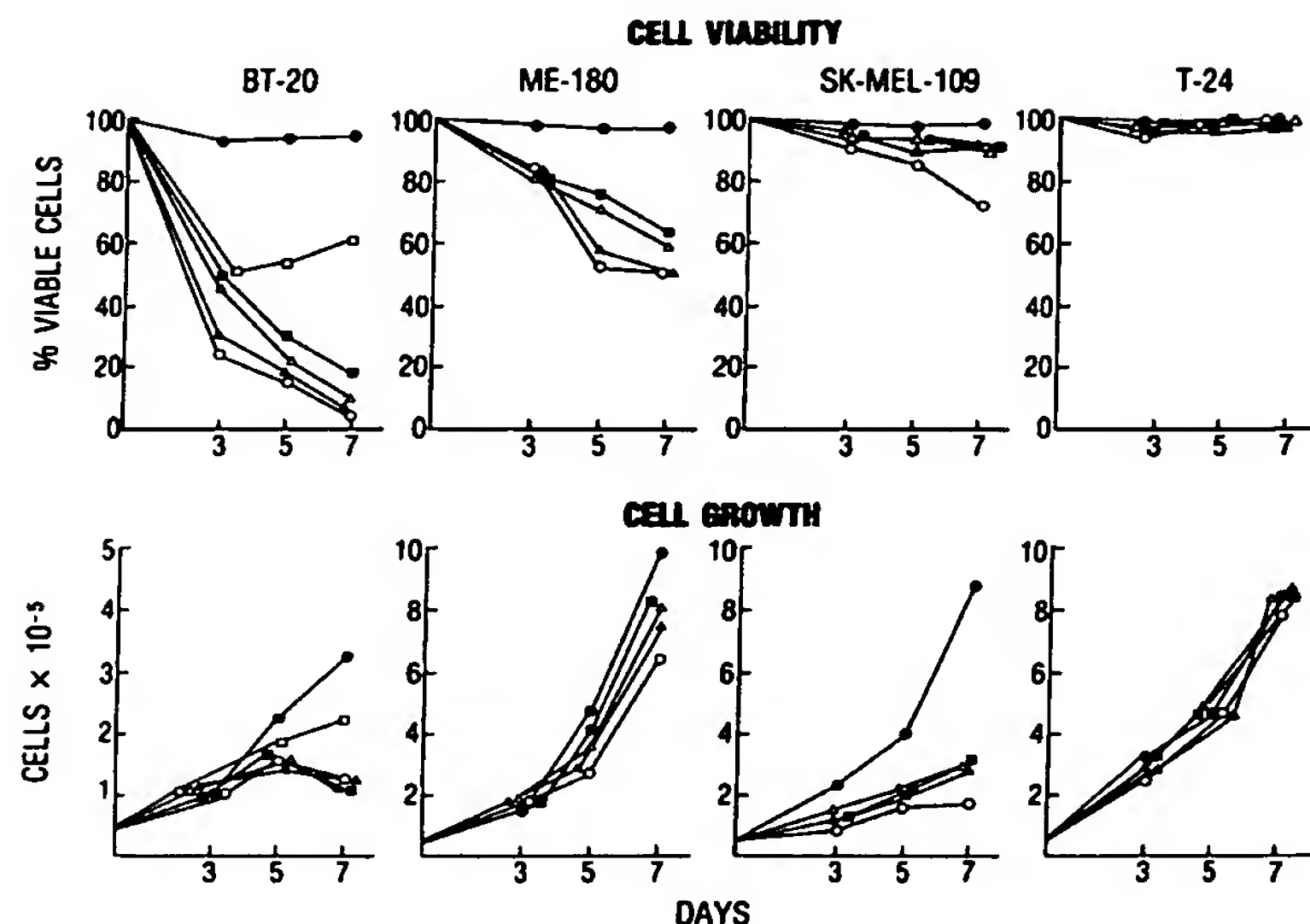


FIG. 1. Effect of hTNF on four human tumor cell lines. hTNF units/ml added to cultures: ●, control cultures; □, 32 units (BT-20); ■, 64 units (ME-180, SK-MEL-109) or 126 units (BT-20, T-24); △, 129 units (ME-180, SK-MEL-109) or 253 units (BT-20, T-24); ▲, 258 units (ME-180, SK-MEL-109) or 505 units (BT-20, T-24); ○, 645 units (ME-180, SK-MEL-109) or 1,263 units (BT-20, T-24).

fractions of culture medium alone. *Limulus* assays indicated no difference in the amount of endotoxin in TNF-active and TNF-inactive (control) fractions.

**Cytotoxic or Cytostatic Effect of hTNF on Human Tumor Cell Lines.** Fig. 1 illustrates the effect of hTNF on four cell lines: BT-20 breast cancer line (cytotoxic effect), ME-180 cervical cancer line (cytotoxic effect), SK-MEL-109 melanoma line (cytostatic effect), and T-24 bladder cancer line (no effect). Table 2 summarizes the results of tests with 23 tumor cell lines and 4 cultures of normal cells. With a 35% or greater reduction in cell viability or cell number at 7 days as the criterion of either cytotoxicity or cytostasis, hTNF has a cytotoxic effect on 7 cell lines, a cytostatic effect on 5 cell lines, and no effect on 15 lines. Three of four cell lines of breast cancer origin were sensitive to the cytotoxic effect of hTNF, whereas the predominant effect of hTNF on melanoma cell lines was cytostasis. None of the four cultures derived from normal tissues was hTNF sensitive.

Table 2. Effect of hTNF on human cell lines

Cytotoxic effect	No effect
SK-MG-4 (astrocytoma)	T-24 (bladder cancer)
MCF-7 (breast cancer)	5637 (bladder cancer)
BT-20 (breast cancer)	MDA-MB-361 (breast cancer)
SK-BR-3 (breast cancer)	S-48 (colon cancer)
ME-180 (cervix cancer)	SK-LC-4 (lung cancer)
SK-CO-1 (colon cancer)	SK-LC-6 (lung cancer)
RPMI 7931 (melanoma)	SK-LC-12 (lung cancer)
	SK-MEL-19 (melanoma)
Cytostatic effect	SK-UT-1 (uterus cancer)
SK-LU-1 (lung cancer)	SAOS-2 (osteogenic sarcoma)
RPMI 4445 (melanoma)	U2OS (osteogenic sarcoma)
SK-MEL-29 (melanoma)	WI-38 (fetal lung)
SK-MEL-109 (melanoma)	MY (normal kidney epithelium)
SK-OV-3 (ovary cancer)	F-136-35-56 (fetal lung)
	F-136-35-56 (fetal skin)

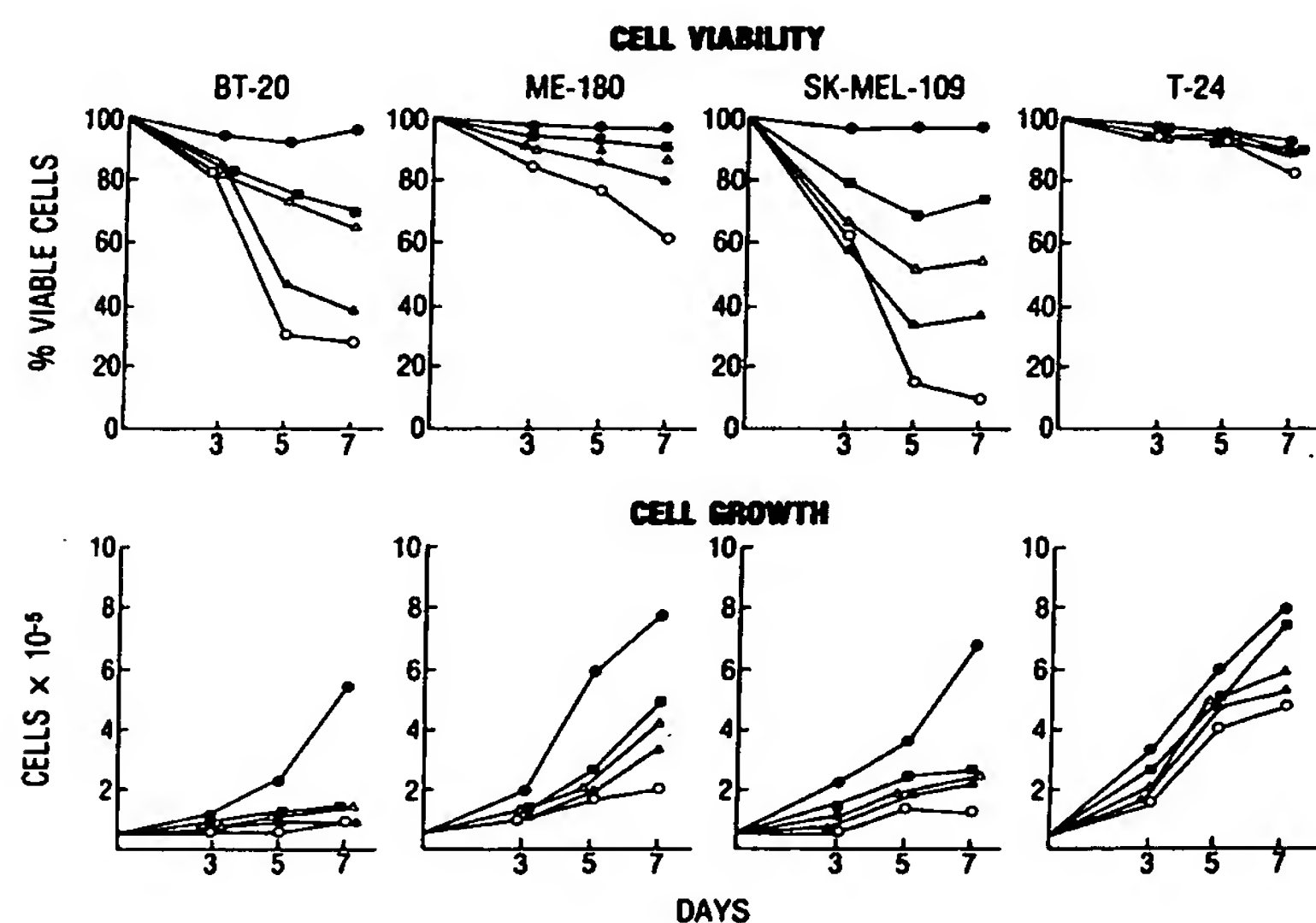


FIG. 2. Effect of recombinant IFN- $\alpha$  on four human tumor cell lines. IFN units/ml added to cultures: ●, control cultures; ■, 781 units; ▲, 3,125 units; ▲, 12,500 units; ○, 50,000 units.

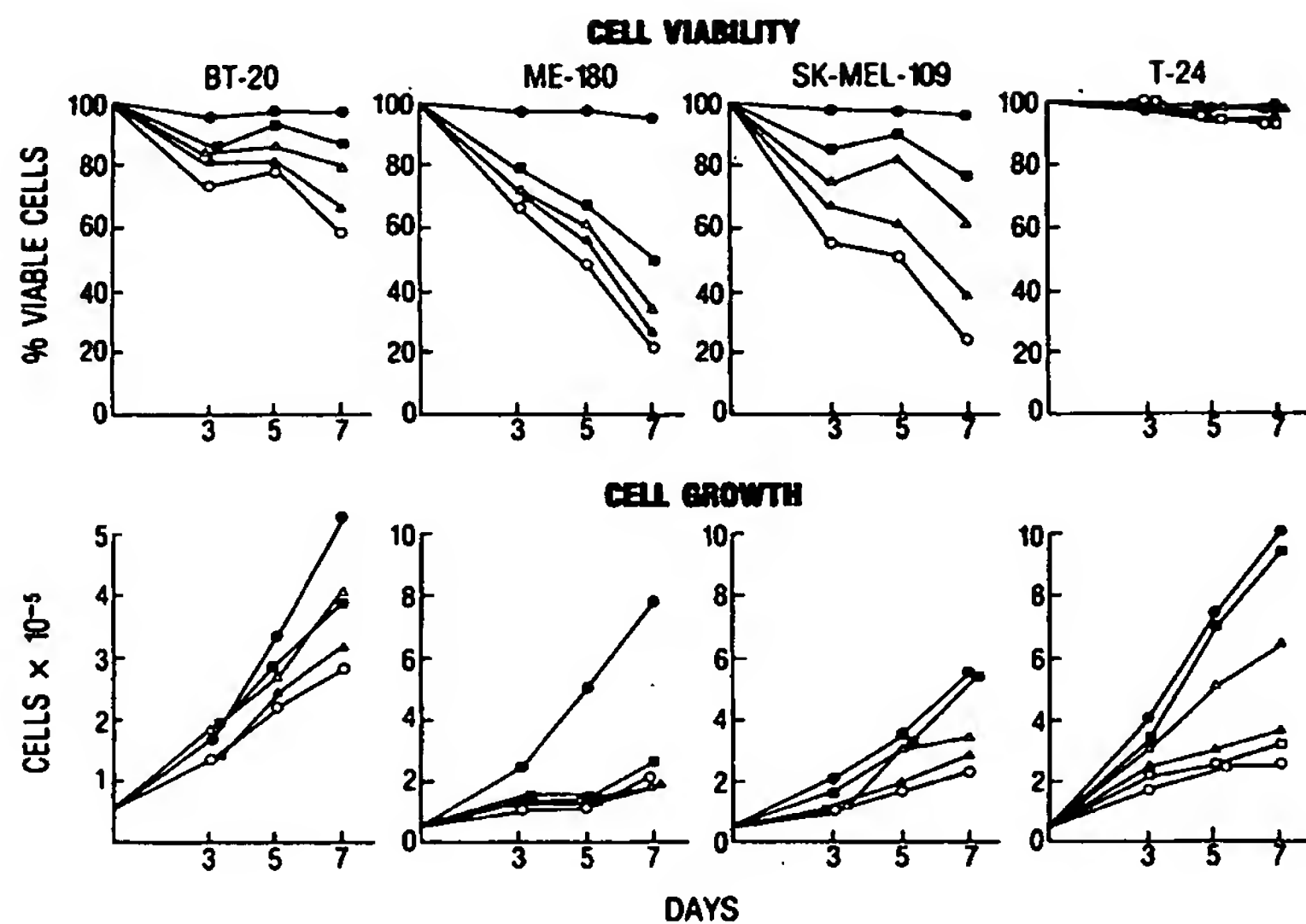


FIG. 3. Effect of IFN- $\gamma$  on four human tumor cell lines. IFN units/ml added to cultures: ●, control cultures; ■, 7.8 units; △, 31.25 units; ▲, 125 units; □, 500 units; ○, 2,000 units.

**Effect of Human IFNs on Human Tumor Cell Lines—Synergistic Action of TNF and Interferon.** Recombinant IFN- $\alpha$  and natural IFNs  $\alpha$ ,  $\beta$ , or  $\gamma$  were tested on a panel of human cell lines. Figs. 2 and 3 show the response of BT-20, ME-180, SK-MEL-109, and T-24 to recombinant IFN- $\alpha$  and natural IFN- $\gamma$  and Table 3 summarizes the results of tests with the various IFN preparations at 7 days. In general, the number of antiviral units of IFN required to cause a 35% or greater cytotoxic effect was higher with IFN- $\alpha$  than with IFN- $\beta$  or - $\gamma$ . None of the IFN preparations showed cytotoxicity or cytostasis on  $L_{45}$  cells, indicating absence of TNF-like activity. Fig. 4 shows the influence of combined treatment with hTNF and recombinant IFN- $\alpha$  or natural IFN- $\gamma$ . Synergistic effects of hTNF and IFN are clearly evident in tests with BT-20, ME-180, and SK-MEL-109, in which the combined cytotoxic effect is greater than that seen with hTNF or IFN alone. Similar results were obtained with mouse TNF and human IFNs  $\alpha$  and  $\gamma$ . By using the approach suggested by Clarke (17), the combined cytotoxic effect of IFN and hTNF was clearly shown to be synergistic rather than additive.

### DISCUSSION

The finding that human cell lines of B-cell origin release a TNF-like molecule after stimulation with PMA suggests that normal B cells have this capacity as well. Whether other human cell types produce TNF remains to be determined, but the fact that cloned lines of mouse histiocytoma do so after endotoxin stimulation indicates the need for further studies of this cell type in humans. The identification of the cytotoxic factor released by the human B-cell lines as the human homolog of mouse TNF

rests on its differential reactivity with TNF-sensitive and TNF-resistant mouse L cells and the similar pattern of response (cytotoxic/cytostatic/no effect) of the human cell panel shown in Table 1 to the human B-cell factor and to mouse TNF (unpublished data). Further facts relating mouse TNF and hTNF are (i) L cells made resistant to hTNF show a cross-resistance to mouse TNF, and (ii) hTNF preparations cause hemorrhagic necrosis of Meth A sarcoma, the standard *in vivo* TNF assay. Although TNF appears to lack species specificity with regard to its antitumor effects, we have found that hTNF has a higher specific activity on human cells as compared to mouse cells. Other features that are shared by mouse and human TNF are acid lability, relative resistance to heat, and augmented toxicity in the presence of metabolic inhibitors such as actinomycin.

Sensitivity to TNF is not an uncommon trait of human cancer cell lines. The normal cell types that have been studied to date have not shown sensitivity to mouse TNF (18) or hTNF, but these tests must be expanded before the apparent tumor specificity of TNF can be assessed. The response of TNF-sensitive tumor cells can be classified either as cytotoxicity or cytostasis. Whether this represents a continuum in response (with cytotoxicity indicating higher sensitivity) or a qualitatively different response is unclear. In studies on the influence of TNF on the cell cycle in sensitive L cells, the most evident effects are an initial accumulation of cells in G2, followed by lysis of cells in telophase (Z. Darzynkiewicz, personal communication). Similar studies with human cell lines showing a predominant cytotoxic or cytostatic effect may clarify the relationship between these two TNF responses.

Neither mouse nor human partially purified TNF has de-

Table 3. Cytotoxic and cytostatic effects of human IFNs on five human cell lines

IFN	Number of anti-viral units causing 35% cytotoxicity or cytostasis at day 7									
	BT-20		ME-180		SK-MEL-109		T-24		WI-38	
	CT	CS	CT	CS	CT	CS	CT	CS	CT	CS
Natural IFN- $\alpha$	270	25	6,500	110	1,250	26	>50,000	1,900	>50,000	230
Recombinant IFN- $\alpha$	3,125	<781	44,000	<781	1,400	<781	>50,000	16,000	>50,000	>50,000
Natural IFN- $\beta$	<49	<49	260	<49	<49	<49	>12,500	<49	>25,000	160
Natural IFN- $\gamma$	130	72	<7.8	<7.8	22	27	>2,000	25	>2,000	>2,000

CT, cytotoxic; CS, cytostatic.



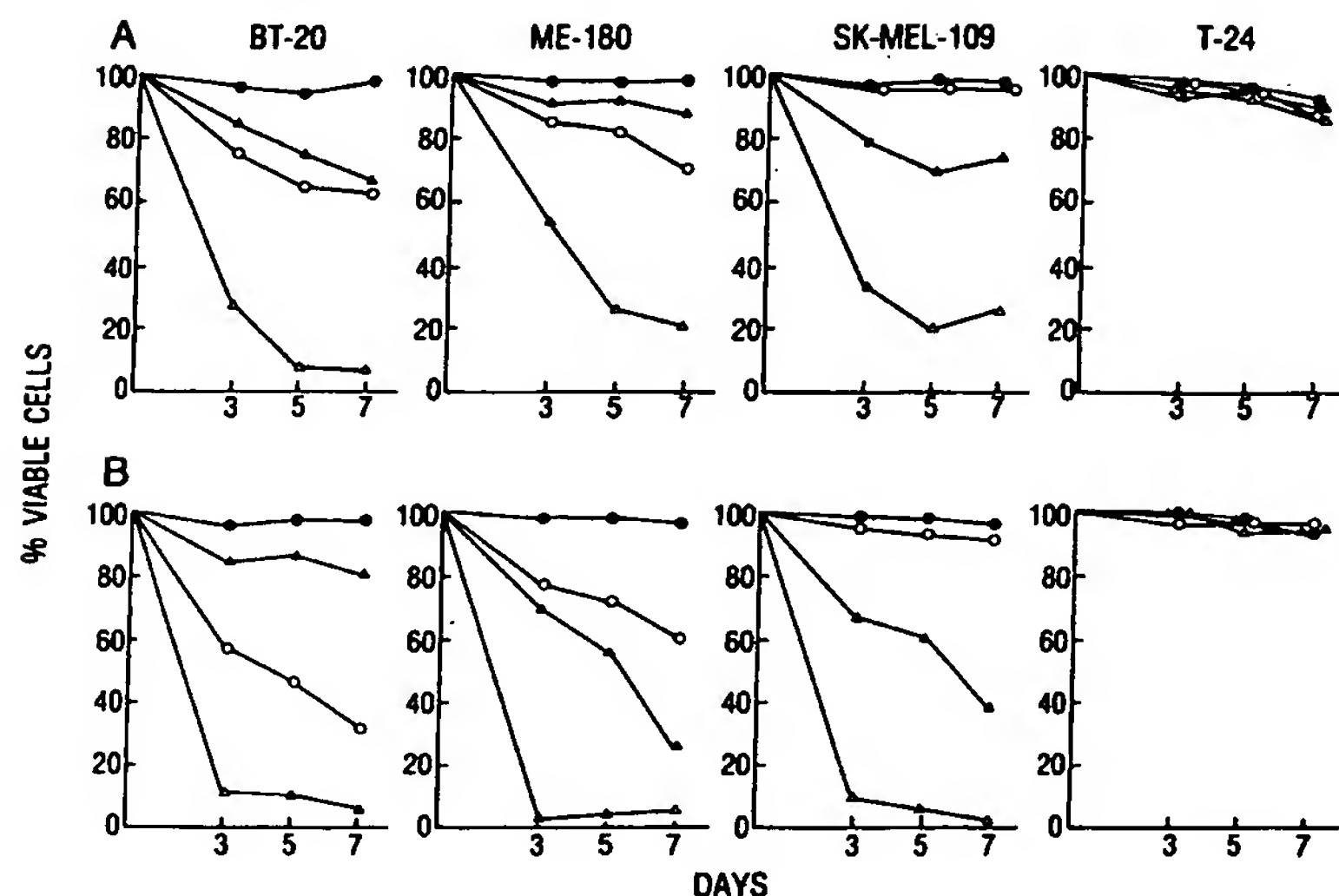


FIG. 4. Effect of hTNF and IFNs on four human tumor cell lines. (A) Recombinant IFN- $\alpha$ . hTNF units/ml and IFN units/ml added to cultures: ●, control cultures; ○, hTNF alone at 4 units (BT-20), 16 units (ME-180), or 33 units (SK-MEL-109, T-24); ▲, IFN alone at 781 units (SK-MEL-109, T-24) or 3,125 units (BT-20, ME-180); △, hTNF and IFN together, in the amounts given above for each cell line. (B) Natural IFN- $\gamma$ . hTNF units/ml and IFN units/ml added to cultures: ●, control cultures; ○, hTNF alone at 25 units (BT-20), 50 units (ME-180), 100 units (SK-MEL-109), or 200 units (T-24); ▲, IFN alone at 31.25 units (BT-20) or 125 units (ME-180, SK-MEL-109, T-24); △, hTNF and IFN together, in the amounts given above for each cell line.

tectable IFN activity, and the various human IFNs lack detectable TNF activity. Thus, TNF and IFN represent two classes of antitumor substances produced by cells of the hematopoietic system and, in the case of IFN- $\beta$ , by other cell types as well. There are a number of other lymphokines that have been described, but these appear to act predominantly by regulating immune responses rather than having a direct cytotoxic or cytostatic effect on cells. Lymphotoxin is an exception in this regard (19, 20), and further studies on the reactivity of lymphotoxin with human cancer cells and its relationship to TNF are required.

In view of the strong anticellular activity of IFNs on a broad range of cancer cells *in vitro*, as shown in this study and in studies by others (21–24), it is surprising that the clinical effects of IFN have not been more evident. The *in vitro* synergistic antitumor action of TNF and IFN observed in the present study suggests that combined treatment with both factors might result in better clinical responses and may also provide clues with regard to requirements for optimal antitumor effects of interferon *in vivo*. If other factors, such as TNF, are necessary for maximal IFN action, deficiencies of these factors in the tumor or in the patient could account for the relatively modest antitumor effects of IFNs observed to date.

This work was supported in part by Grants CA-08748 and AI-17920 from the National Institutes of Health and by a grant from the Cancer Research Institute.

1. Carswell, E. A., Old, L. J., Kassel, R. L., Green, S., Fiore, N. & Williamson, B. (1975) *Proc. Natl. Acad. Sci. USA* 72, 3666–3670.
2. Helson, L., Green, S., Carswell, E. & Old, L. J. (1975) *Nature (London)* 258, 731–732.
3. Mannel, D. N., Meltzer, M. S. & Mergenhagen, S. E. (1980) *Infect. Immun.* 28, 204–211.

4. Kull, F. C. & Cuatrecasas, P. (1981) *J. Immunol.* 126, 1279–1283.
5. Haranaka, K. & Satomi, N. (1981) *Jpn. J. Exp. Med.* 51, 191–194.
6. Green, S., Dobrjansky, A., Chiasson, M. A., Carswell, E., Schwartz, M. K. & Old, L. J. (1977) *J. Natl. Cancer Inst.* 59, 1519–1522.
7. Matthews, N. & Watkins, J. F. (1978) *Br. J. Cancer* 38, 302–309.
8. Ostrove, J. M. & Gifford, G. E. (1979) *Proc. Soc. Exp. Biol. Med.* 160, 354–358.
9. Green, S., Dobrjansky, A., Carswell, E. A., Kassel, R. L., Old, L. J., Fiore, N. & Schwartz, M. K. (1976) *Proc. Natl. Acad. Sci. USA* 73, 381–385.
10. Matthews, N., Ryley, H. C. & Neale, M. L. (1980) *Br. J. Cancer* 42, 416–422.
11. Ruff, M. R. & Gifford, G. E. (1980) *J. Immunol.* 125, 1671–1677.
12. Matthews, N. (1978) *Br. J. Cancer* 38, 310–315.
13. Mannel, D. N., Moore, R. N. & Mergenhagen, S. E. (1980) *Infect. Immun.* 30, 523–530.
14. Pickering, L. A., Kronenberg, L. H. & Stewart, W. E., II (1980) *Proc. Natl. Acad. Sci. USA* 77, 5938–5942.
15. Houghton, A. N., Taormina, M. C., Ikeda, H., Watanabe, T., Oettgen, H. F. & Old, L. J. (1980) *Proc. Natl. Acad. Sci. USA* 77, 4260–4264.
16. Stewart, W. E., II (1979) *The Interferon System* (Springer, Vienna).
17. Clarke, D. A. (1958) *Ann. N.Y. Acad. Sci.* 76, 915–921.
18. Old, L. J. (1976) *Clin. Bull.* 6, 118–120.
19. Granger, G. A., Yamamoto, R. S., Fair, D. S. & Hiserodt, J. C. (1978) *Cell Immunol.* 38, 388–402.
20. Rosenau, W. (1981) *Int. J. Immunopharmacol.* 3, 1–8.
21. Blalock, J. E., Georgiades, J. A., Langford, M. P. & Johnson, H. M. (1980) *Cell Immunol.* 49, 390–394.
22. DeLey, M., van Damme, J., Claeys, H., Weening, H., Heine, J. W., Billiau, A., Vermeylen, C. & De Somer, P. (1980) *Eur. J. Immunol.* 10, 877–883.
23. Rubin, B. Y. & Gupta, S. L. (1980) *Proc. Natl. Acad. Sci. USA* 77, 5928–5932.
24. Borden, E. C., Hogan, T. F. & Voelkel, J. G. (1982) *Cancer Res.* 42, 4948–4953.



## ROLE OF INTERLEUKIN 1 IN ANTI-IMMUNOGLOBULIN-INDUCED B CELL PROLIFERATION\*

BY MAUREEN HOWARD, STEVEN B. MIZEL, LAWRENCE LACHMAN,  
JOHN ANSEL, BARBARA JOHNSON, AND WILLIAM E. PAUL

*From the Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20205, The Microbiology Program, Pennsylvania State University, University Park, Pennsylvania 16802, and Immunex Corporation, 51 University Building, Seattle, Washington 98101*

The use of anti-Ig antibodies to stimulate polyclonal proliferation through Ig receptors expressed on the B cell membrane has been a popular model for antigen-driven B cell activation (1-7). Initially it was proposed that anti-Ig-induced proliferation proceeded independently of T cells and other accessory cells, as vigorous depletion of such cell types failed to prevent activation (3, 8). However, certain features of anti-Ig-induced B cell activation challenged the view that the response was a simple consequence of the interaction of anti-Ig and membrane Ig. In particular, the relationship between cell density and magnitude of the proliferative response was nonlinear and rapidly declined to background proliferation levels at cell numbers below  $10^5$  per microtiter well (3, 6). Furthermore, the optimum doses of anti-Ig used in such studies far exceeded the amount required to saturate membrane Ig receptors and induce capping of membrane components. Precise delineation of the stimuli required for anti-Ig-induced B cell activation requires a functional assay in which contaminating accessory cells potentially capable of endogenous factor production have been excluded. To this end, we have investigated conditions required for polyclonal activation of highly purified mouse B lymphocytes cultured at low cell density (e.g.,  $5 \times 10^4$  cells/well). Using such an assay, we have shown the role of a T cell-derived factor, designated B cell growth factor (BCGF),<sup>1</sup> in anti-IgM-induced B cell proliferation. (9). In this report we demonstrate that in the presence of optimum amounts of BCGF, anti-IgM-induced proliferation of B cells cultured at lower densities ( $1-2 \times 10^4$  cells/well) is enhanced by the addition of a second soluble factor. Biochemical analyses identify this second cofactor as the previously described monokine interleukin 1 (IL-1).

### Materials and Methods

*Mice.* BALB/cJ mice were obtained from The Jackson Laboratory, Bar Harbor, ME, and used at 8-12 wk of age. CH3/HeJ mice, also obtained from The Jackson Laboratory, were used at 5-8 wk of age.

\* Supported in part by grant AI 17559 from the U. S. Public Health Service and grant PCM-8110370 from the National Science Foundation, both to S. B. Mizel.

<sup>1</sup> Abbreviations used in this paper: BCGF, B cell growth factor; IEF, isoelectricfocusing; IL-1, IL-2, interleukins 1 and 2; PAGE, polyacrylamide gel electrophoresis; PMA, phorbol myristate acetate; UV, ultraviolet.

**Anti-IgM Antibodies.** Affinity-purified goat anti-mouse Ig specific for  $\mu$  heavy chains (anti-IgM) was prepared as described previously (6).

**Culture Medium.** The culture medium used throughout was RPMI 1640 (Gibco Laboratories, Grand Island, NY) supplemented with 10% fetal calf serum (Reheis, Kankakee, IL), penicillin (50  $\mu$ g/ml), streptomycin (50  $\mu$ g/ml), gentamycin (100  $\mu$ g/ml), L-glutamine (200 mM), and 2-mercaptoethanol ( $5 \times 10^{-6}$  M) (2ME).

**Growth Factor Preparations.** EL4 supernatant was produced by stimulating a cloned subline of EL4 thymoma with 10 ng/ml phorbol myristate acetate (PMA) as described elsewhere (10). Cell-free supernatants were collected after 48 h and depleted of PMA by adsorption on activated charcoal (11). For some experiments, BCGF was partially purified from EL4 supernatants by phenylsepharose chromatography or isoelectricfocusing (IEF) (12).<sup>2</sup>

Macrophages were stimulated to produce growth factors by four different procedures.

(a) Stimulation of the cloned murine monocytic cell line P388D<sub>1</sub> with 1  $\mu$ g/ml PMA, as described by Mizel et al. (13). Cell-free supernatants were collected after 5 d and depleted of PMA by two successive adsorptions on activated charcoal. For some experiments, the supernatants were concentrated ~200-fold by Amicon filtration, then fractionated by gel filtration on an AcA54 column as described previously (10) to yield material of 10,000–20,000 mol wt.

(b) Treatment of P388D<sub>1</sub> cells with 2–4 min ultraviolet (UV) irradiation (200–400 mJ/cm<sup>2</sup>) using a bank of four Westinghouse FS20 bulbs followed by their culture for 24 h in serum-free RPMI as described (Ansel, J., T. Luger, and I. Green, manuscript submitted for publication).

(c) A superinduction protocol, designed to enhance IL-1 production by P388D<sub>1</sub> cells for factor purification purposes. This protocol involved co-culture of cells with protein and RNA synthesis inhibitors in addition to PMA for 5 h, removal of these agents by washing, then continued culture of the activated P388D<sub>1</sub> cells for a further 24 h (14).

(d) Stimulation of human acute monocytic leukemia cells with endotoxin and collection of cell-free supernatants 48 h later, as described by Lachman et al. (15).

**B Cell Co-stimulator Assay.** Full details of this assay are given elsewhere (9). Briefly, splenic B cells were purified by the procedure of Leibson et al. (16), then cultured at densities ranging from  $10^3$  to  $5 \times 10^5$ /well in 200  $\mu$ l medium in flat-bottomed 96-well microtiter plates (0.32 cm<sup>2</sup> growth area). Some cultures contained affinity-purified goat anti-IgM antibody at 5–10  $\mu$ g/ml, and/or dilutions of the various growth factor preparations. Cultures were incubated at 37°C in a humidified atmosphere of air containing 7.5% CO<sub>2</sub> for 72 h. The proliferative response of these cultures was determined by adding [<sup>3</sup>H]thymidine (1  $\mu$ Ci, 6.7 Ci/mmol; New England Nuclear, Boston, MA) for the last 12–16 h of culture and measuring [<sup>3</sup>H]thymidine incorporation.

**Purification of IL-1.** Mouse and human IL-1 were purified according to previously published procedures (14, 17, 18). Mouse IL-1 in super-induced P388D<sub>1</sub> supernatants was purified to apparent homogeneity using the following protocol: gel filtration chromatography, IEF, sulphopropyl Sephadex cationic exchange chromatography, and tris glycinate polyacrylamide gel electrophoresis (PAGE). Human IL-1 was purified to high specific activity using the following procedures in sequence: hollow fiber diafiltration and ultrafiltration, IEF, and tris glycinate PAGE (18, 19). Criteria for purity of the preparations yielded by these protocols are outlined elsewhere (14, 17, 18).

**IL-1 Assay.** Mouse IL-1 was measured as a co-mitogen with phytohemagglutinin in a mouse thymocyte [<sup>3</sup>H]thymidine incorporation assay (19). Human IL-1 was measured as a direct mitogen for mouse thymocytes in a [<sup>3</sup>H]thymidine incorporation assay (18).

## Results

**BCGF Dependence of Anti-Ig-induced B Cell Proliferation.** As has been shown previously (3, 6), a careful analysis of the relationship between B cell density and anti-IgM-induced proliferation revealed a sharp decline in response at cell numbers below  $10^5$  per microtiter well (Fig. 1 A). This result strongly suggested that in addition to anti-IgM and the B lymphocytes, other elements in the culture were limiting. We have

<sup>2</sup> Farrar, J., M. Howard, J. Fuller-Farrar, and W. E. Paul. Biochemical characterization of a murine B cell growth factor distinct from interleukin 2. Manuscript submitted for publication.

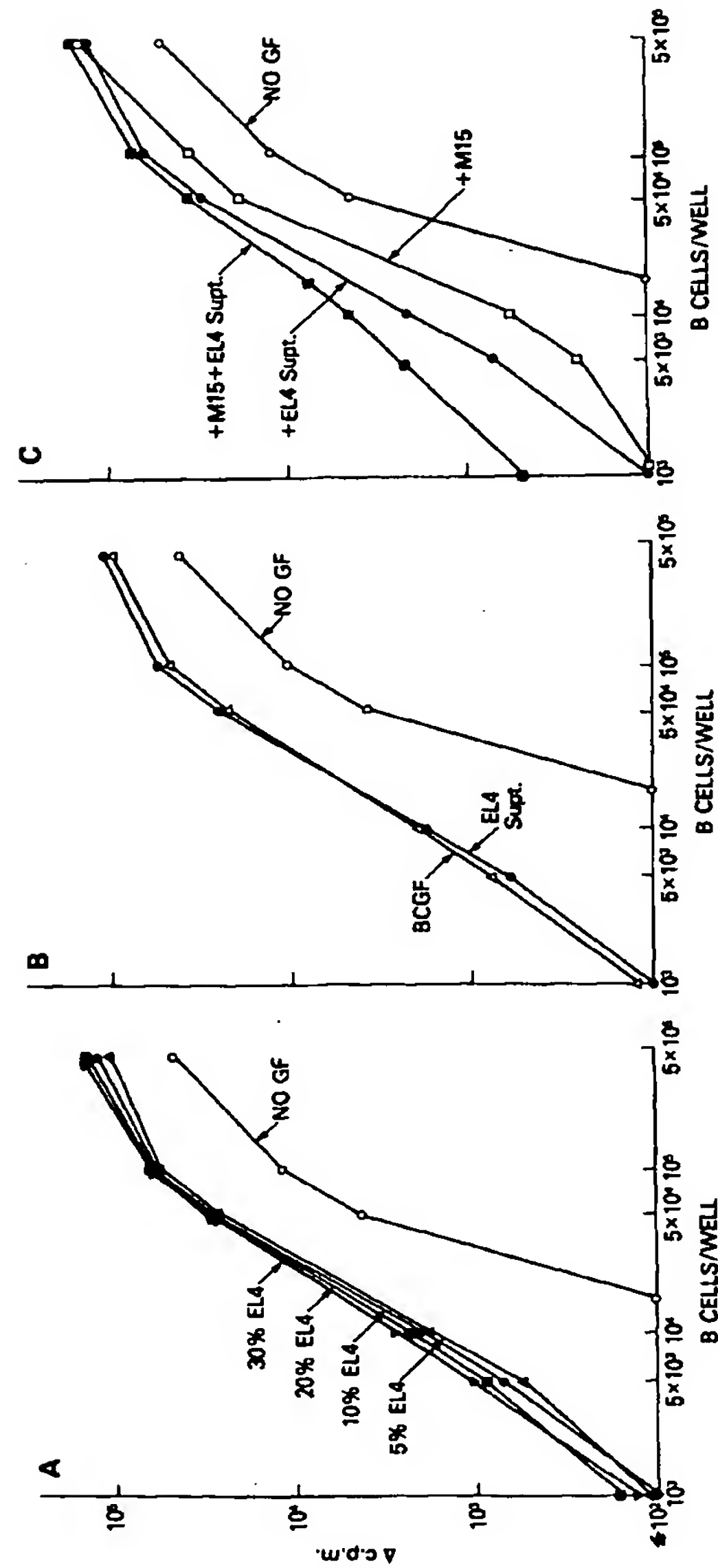


Fig. 1 Relationship between cell density and proliferative response of purified B cells cultured with anti-IgM (5  $\mu$ g/ml) alone (NO GF) or with various preparations of T cell or macrophage-derived factors. Cells were cultured for ~2.5 d, with the addition of [ $^3$ H]thymidine for the final 16 h. Results reflect the difference in [ $^3$ H] thymidine incorporation obtained in cultures containing anti-IgM versus those lacking anti-IgM. All results represent means of duplicate or triplicate cultures. (A) Cultures were supplemented with EL4 supernatant added at final concentrations ranging from 5 to 30%. (B) Cultures were supplemented either with 10% EL4 supernatant ( $\bullet$ ), or with an IEF-purified preparation of BCGF ( $\Delta$ ) added to cultures at a 1:40 dilution, a concentration found to be saturating for co-stimulator activity at the B cell density of  $5 \times 10^4$ /well. (C) Cultures were supplemented with 10% EL4 supernatant ( $\bullet$ ), M15 added at final dilution of 1:16 ( $\square$ ), or both EL4 supernatant and M15 ( $\blacksquare$ ). M15 was purified by gel filtration, and corresponds to the hatched area shown in Fig. 2.

already shown (9) that induced supernatants from the mouse thymoma EL4 produce two striking effects in cultures of anti-IgM stimulated B cells: (a) a shift in the antibody concentration/response relation such that 5  $\mu\text{g/ml}$  anti-IgM plus EL4 supernatant induces levels of proliferation greater than those induced by 50  $\mu\text{g/ml}$  anti-IgM alone, and (b) a shift in the cell density/response relation such that a very substantial response to 5–50  $\mu\text{g/ml}$  anti-IgM can be obtained with as few as  $5 \times 10^4$  B cells/well. Here we see that EL4 supernatant enhanced the B cell proliferative response at cell densities ranging from  $5 \times 10^3/\text{well}$  to  $5 \times 10^5/\text{well}$ , and that essentially identical results were obtained with concentrations of EL4 supernatant ranging between 5 and 30% indicating saturation conditions (Fig. 1 A). The B cell co-stimulating factor in EL4 supernatants has been examined in detail elsewhere (9, 12).<sup>2</sup> These studies show that the factor, designated BCGF, has an apparent molecular weight by gel filtration of 18,000 and is clearly distinct from interleukin 2 (IL-2) by gel filtration, phenylsepharose chromatography, sodium dodecyl sulfate-PAGE, and IEF, as well as by cellular absorption studies. BCGF partially purified by IEF caused the same maximal enhancement of response to anti-IgM as did unfractionated EL4 supernatant (Fig. 1 B). This strongly suggests that the enhancing activity of EL4 is attributable to its BCGF content.

*A Monokine Enhances BCGF-dependent Anti-IgM-induced B Cell Proliferation.* To explore the possibility of a macrophage-derived factor also being involved in anti-IgM-induced B cell proliferation, we selected B cell co-stimulator assay conditions that yielded low proliferative responses on a per cell basis to anti-IgM even in the presence of saturating amounts of EL4 supernatant, i.e.,  $1-2 \times 10^4$  B cells/microtiter well plus 10% EL4 supernatant (see Fig. 1 A), and examined the effects of supplementing these cultures with induced supernatants from the cloned murine macrophage cell line P388D<sub>1</sub>. Others have shown that such supernatants are excellent sources of monokines such as IL-1 (13, 14, 20). Unfractionated PMA-induced P388D<sub>1</sub> supernatants were found to profoundly inhibit both BCGF-dependent anti-IgM-induced and unstimulated background levels of B cell proliferation (Table I). As unfractionated supernatants may contain both inhibitors and enhancing factors, we repeated these experiments using fractionated components of induced P388D<sub>1</sub> supernatants. Thus, P388D<sub>1</sub> supernatants were fractionated by gel filtration using a calibrated AcA54 column exactly as described elsewhere (10). Each column fraction was dialyzed against culture medium, and added at five serial dilutions to cultures containing  $10^4$  highly purified B cells, anti-IgM antibodies, and a saturating amount of the EL4 supernatant. The results obtained at a 1:16 dilution of column fractions revealed that proliferation of

TABLE I  
*Effect of Unfractionated PMA-induced Macrophage Supernatants on B Cell Proliferation*

Macrophage factors	T cell factors	cpm/ $5 \times 10^4$ B cells	
		No anti-IgM	Plus Anti-IgM
—	—	1,713 $\pm$ 267	2,133 $\pm$ 301
P388D <sub>1</sub> supernatant	—	10 $\pm$ 41	189 $\pm$ 34
—	EL4 supernatant	3,436 $\pm$ 1,021	16,277 $\pm$ 222
P388D <sub>1</sub> supernatant	EL4 supernatant	358 $\pm$ 27	482 $\pm$ 103

25% P388D<sub>1</sub> supernatant; 10% EL4 supernatant; 5  $\mu\text{g/ml}$  anti-IgM.



such low numbers of B cells was markedly enhanced by the addition of macrophage-derived product(s) in the molecular weight range of 10–20,000 (Fig. 2). This material was not mitogenic for resting B cells, nor did it stimulate resting B cells in the presence of EL4 supernatant only. For convenience, the material was initially termed M15 to signify its origin (i.e., macrophage) and approximate molecular weight (i.e., 15,000). A pool was made of those fractions constituting the major peak of M15 activity (see hatched area of Fig. 2), and this pooled material was titrated under identical assay conditions. The results showed M15 to be (a) inhibitory at high concentrations and (b) active over a wide concentration range below the inhibitory doses (Fig. 3 A). These findings have been reproduced using several different batches of M15 prepared from separate PMA-induced P388D<sub>1</sub> supernatants. Similar B cell co-stimulator activity was found in supernatants obtained from UV-irradiated P388D<sub>1</sub> cells (Fig. 3 B), thereby allaying concern regarding the potential involvement of PMA in the co-stimulatory activity of P388D<sub>1</sub> supernatants.

The effect of M15 on the cell density/response relationship of anti-IgM-induced B cell proliferation was tested using a 1:16 dilution of M15, i.e., the optimal concentration of this material (see Fig. 3 A). B cell proliferation was enhanced at all cell densities tested (Fig. 1 C). When the same experiment was performed in the presence of a saturating amount of the T cell-derived lymphokine BCGF, a synergistic effect of the two factors was observed, particularly at low densities (Fig. 1 C). Identical results were obtained when a PMA-free source of M15 (i.e., supernatant of UV-irradiated P388D<sub>1</sub> cells) was used (Fig. 4 A). The results shown in Fig. 1 have been corrected for proliferation obtained in the absence of anti-IgM antibodies, and thus represent anti-IgM-dependent activation. We emphasize that both M15 (see Figs. 2 and 3) and

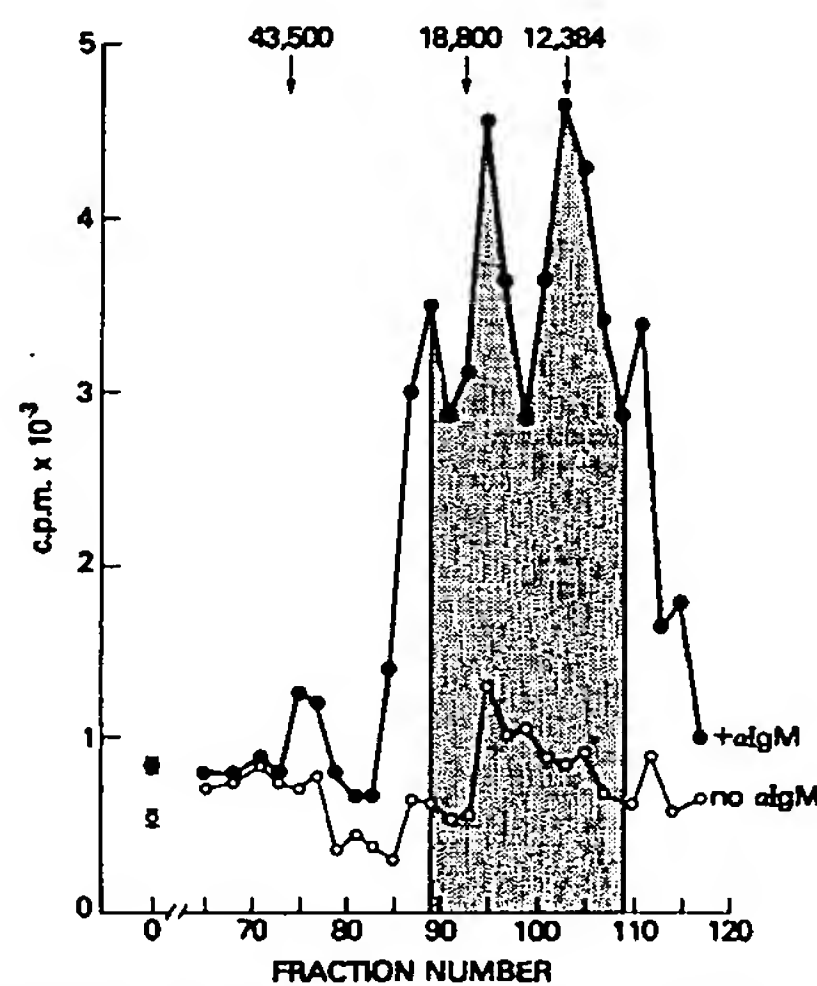


FIG. 2. Gel filtration analysis of B cell co-stimulator activity in PMA-induced P388D<sub>1</sub> supernatant. Fractions collected from an AcA54 column were dialysed against culture medium, then added at a final concentration of 1:16 to cultures containing  $10^4$  purified B cells per well and 10% EL4 supernatant with or without anti-IgM antibodies (5  $\mu$ g/ml). Proliferation was assessed at day 2.5 by [ $^3$ H]thymidine uptake. The calibrating molecular weight markers were ovalbumin (43,500), myoglobin (18,800), and cytochrome c (12,384). All results represent duplicate cultures.

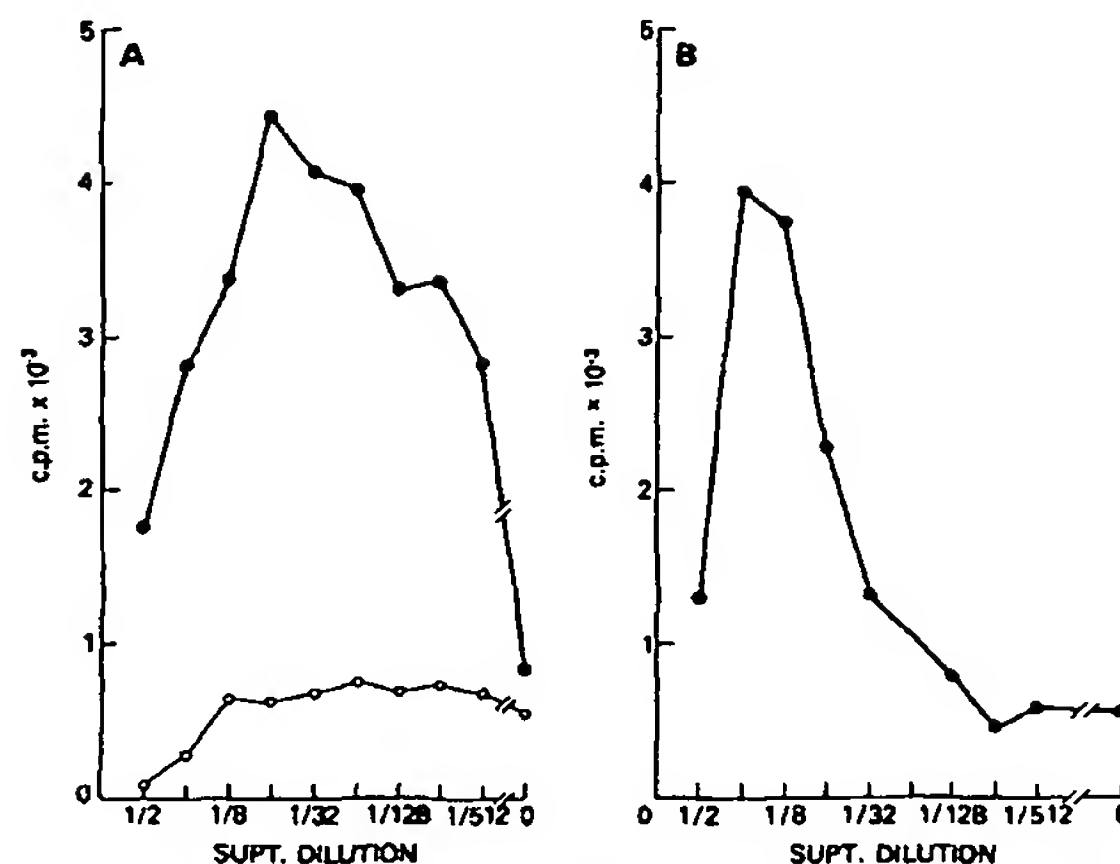


FIG. 3. B cell co-stimulator activity in (A) a 10,000–20,000-mol wt pool made from PMA-induced P388D<sub>1</sub> supernatant (equivalent to hatched area of Fig. 2), or (B) supernatant collected from UV-irradiated P388D<sub>1</sub> cells (refer to Materials and Methods). Supernatants were added at final dilutions of from 1/2 to 1/512 to cultures containing 10<sup>4</sup> purified B cells and 10% EL4 supernatant, with (●) or without (○) anti-IgM antibodies (5 μg/ml). Proliferation was assessed at day 2.5 by [<sup>3</sup>H]-thymidine uptake. Results represent duplicate cultures.

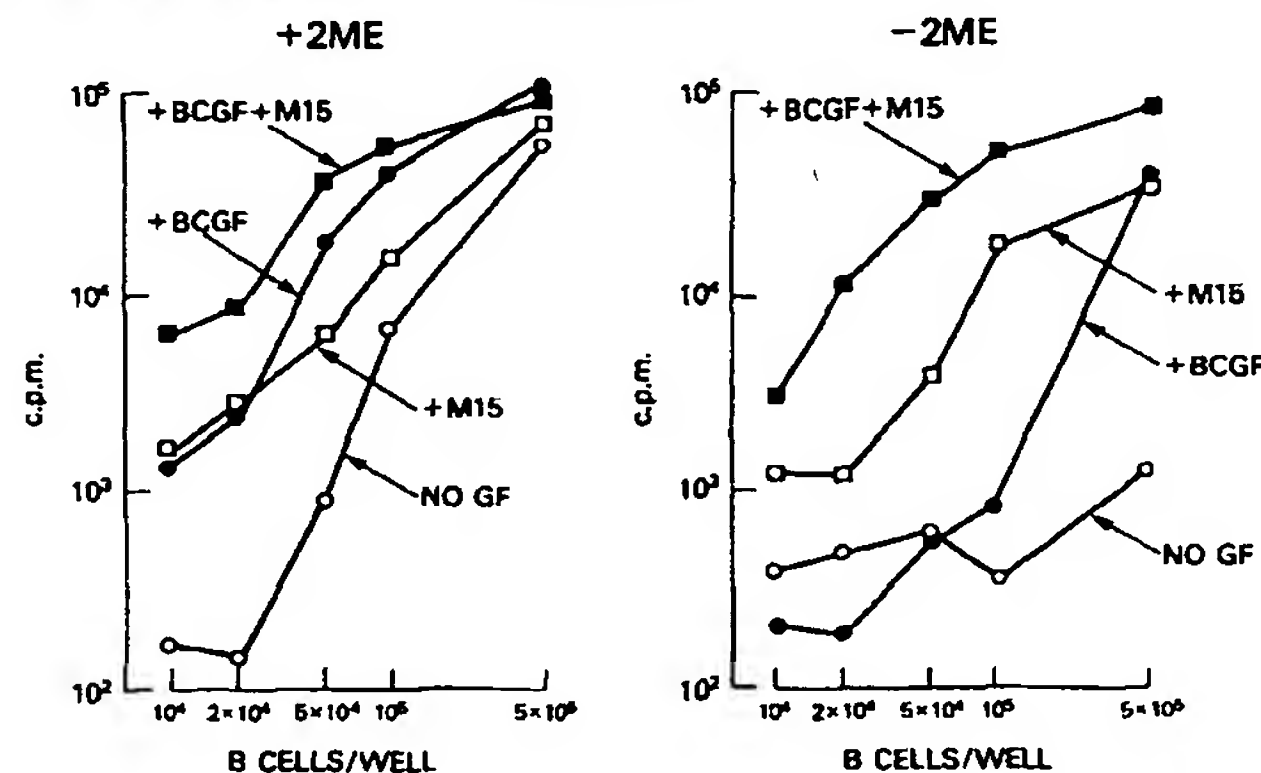


FIG. 4. Proliferative response of purified B cells cultured with anti-IgM (5 μg/ml) together with various co-factors either in culture medium containing  $2 \times 10^{-6}$  M 2ME or in 2ME-free medium. Co-factor supplements were: (a) phenyl Sepharose-purified BCGF, prepared as in reference 12 (also footnote 2) and added at a 1:50 final dilution found to be saturating for co-stimulator activity at the B cell density of  $5 \times 10^4$ /well; (b) supernatant collected from UV-irradiated P388D<sub>1</sub> cells and added at a final dilution of 1:4 (designated M15); (c) medium only (NO GF); (d) a mixture of (a) and (b). Other assay conditions were as in Fig. 1.

BCGF (9) produce little proliferation in the absence of anti-IgM antibodies. In summary, these data strongly imply roles for both BCGF and M15 in anti-IgM-induced B cell proliferation.

**Absolute Need for M15 in 2ME-free Cultures.** M15-mediated enhancement of anti-Ig induced BCGF-dependent B cell proliferation was highly reproducible from experiment to experiment. However, the magnitude of its effect was often relatively small (two- to fourfold), making it difficult to clearly determine whether BCGF and M15

acted synergistically or additively. This problem was resolved by taking advantage of the observation that B cell proliferation experiments performed in culture medium lacking 2ME show an exaggerated need for exogenously added M15. Fig. 4 shows B cell proliferative responses to anti-IgM plus cofactors in medium containing or lacking 2ME. While the maximum proliferation levels obtained in the presence of anti-IgM, BCGF, and M15 were the same in both media, a significant response to anti-IgM alone, i.e., in the absence of exogenously added BCGF and M15, was only obtained when B cells were cultured in 2ME-containing medium; in the absence of 2ME a response to anti-IgM and BCGF was only observed at  $5 \times 10^5$  cells/well, the highest cell density tested (Fig. 4B). In cultures lacking 2ME, M15 has a marked costimulatory effect at all densities. These data prompt speculation regarding the possibility of endogenous 2ME-mediated M15 production by residual macrophages and/or B cells in our standard BCGF co-stimulator assay. The striking M15-dependent enhancement of B cell proliferation observed in 2ME-free cultures clearly establishes the synergy between BCGF and M15, suggesting that both co-factors operate on the same population of B cells.

*Identification of M15 as IL-1.* In terms of approximate molecular weight and cellular origin, M15 resembles the previously described monokine IL-1. To investigate the possible identity of these two factors, we tested murine and human IL-1, purified according to previously established procedures, for M15 activity, i.e., ability to enhance the proliferation of low numbers of purified B cells cultured with anti-IgM and saturating doses of EL4 supernatant.

Murine IL-1 can be obtained in large quantities by a superinduction procedure involving pre-culture of P388D<sub>1</sub> cells with PMA in the presence of protein and RNA inhibitors (14). Mizel and Mizel (14) have purified the IL-1 in this crude supernatant to apparent homogeneity by a modification of a previously published fractionation scheme involving gel filtration, IEF, sulpho-propyl Sephadex cationic exchange, and tris glycinate PAGE. As expected, the 15,000-mol wt material obtained by gel filtration of supernatant from superinduced P388D<sub>1</sub> cells showed excellent M15 activity when added to cultures of highly purified B cells, anti-IgM, and saturating amounts of EL4 supernatant (data not shown). The 15,000-mol wt material was then further purified by IEF, sulpho-propyl Sephadex cationic exchange, and tris glycinate PAGE. Fractions from the tris glycinate gel were dialyzed, then assayed for M15 activity using the B cell co-stimulating assay described above and for IL-1 activity using the conventional mouse thymocyte proliferation assay. The two activity profiles were indistinguishable (Fig. 5), indicating either that M15 is IL-1 or that there is a great biochemical similarity between them.

Human and murine IL-1 are functionally interchangeable in a variety of T cell assays (S. B. Mizel, unpublished observations). Thus, to further assess the relationship between M15 and IL-1, we tested purified human IL-1 for M15 activity. Human IL-1 can be obtained in large quantities by incubating freshly collected acute monocytic leukemia cells with endotoxin (15). Lachman et al. (17) have purified the IL-1 in this crude supernatant to high specific activity by sequential fractionation involving hollow fiber filtration, IEF, and tris glycinate PAGE (17, 18). This combination of procedures generally yields a preparation that is free of endotoxin, as assessed by the Limulus assay (L. Lachman, unpublished data). The 10–50,000-mol wt material obtained by hollow fiber filtration was subjected to IEF, and dialyzed fractions were

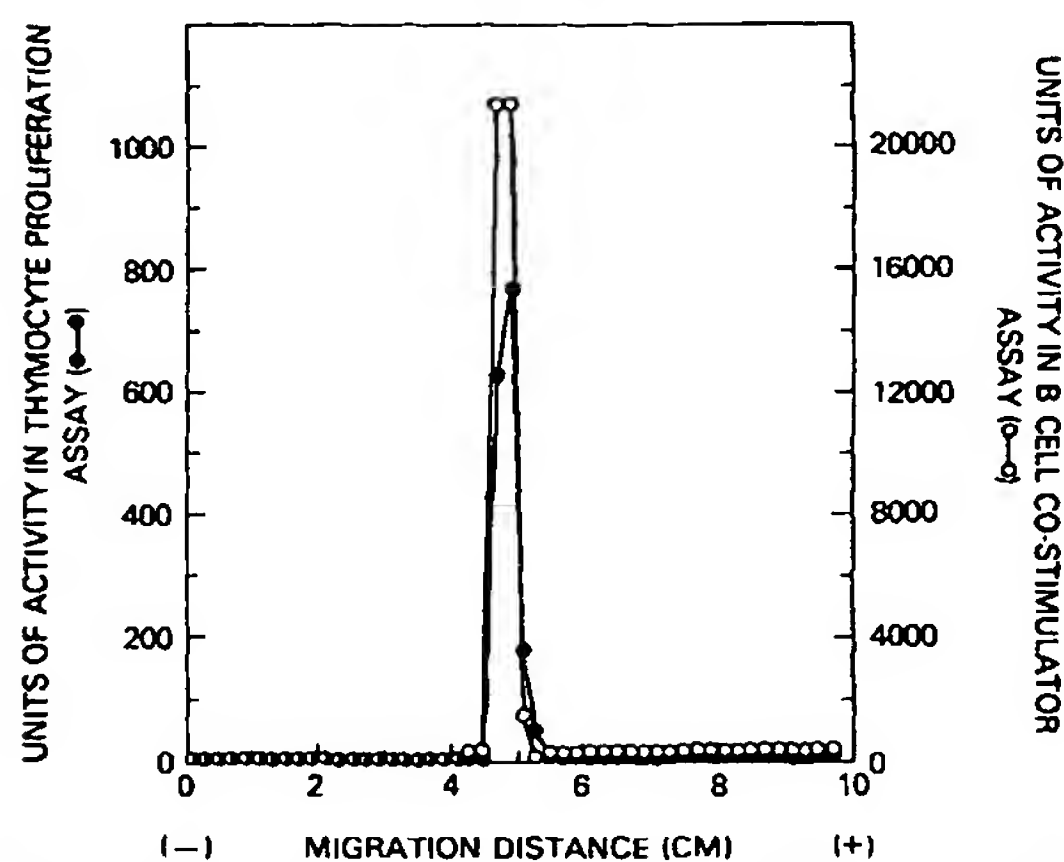


FIG. 5. B cell co-stimulating activity of highly purified murine IL-1. Murine IL-1 from superinduced P388D<sub>1</sub> supernatants was purified to apparent homogeneity as described in text. Dialyzed fractions obtained from the final tris glycinate gel electrophoresis were added at several dilutions ranging from 1/8 to 1/2048 to cultures of  $2 \times 10^4$  B cells, 10% EL4 supernatant, and anti-IgM (5  $\mu$ g/ml) (○), or to cultures of  $1.5 \times 10^6$  mouse thymocytes and phytohemagglutinin (●). In both cases, proliferation was assessed at day 2.5 by [ $^3$ H]thymidine uptake. For each fraction, relative units of activity represent the inverse of the dilution that produced 50% of the maximum proliferation obtained in that assay. All results represent means of duplicate cultures.

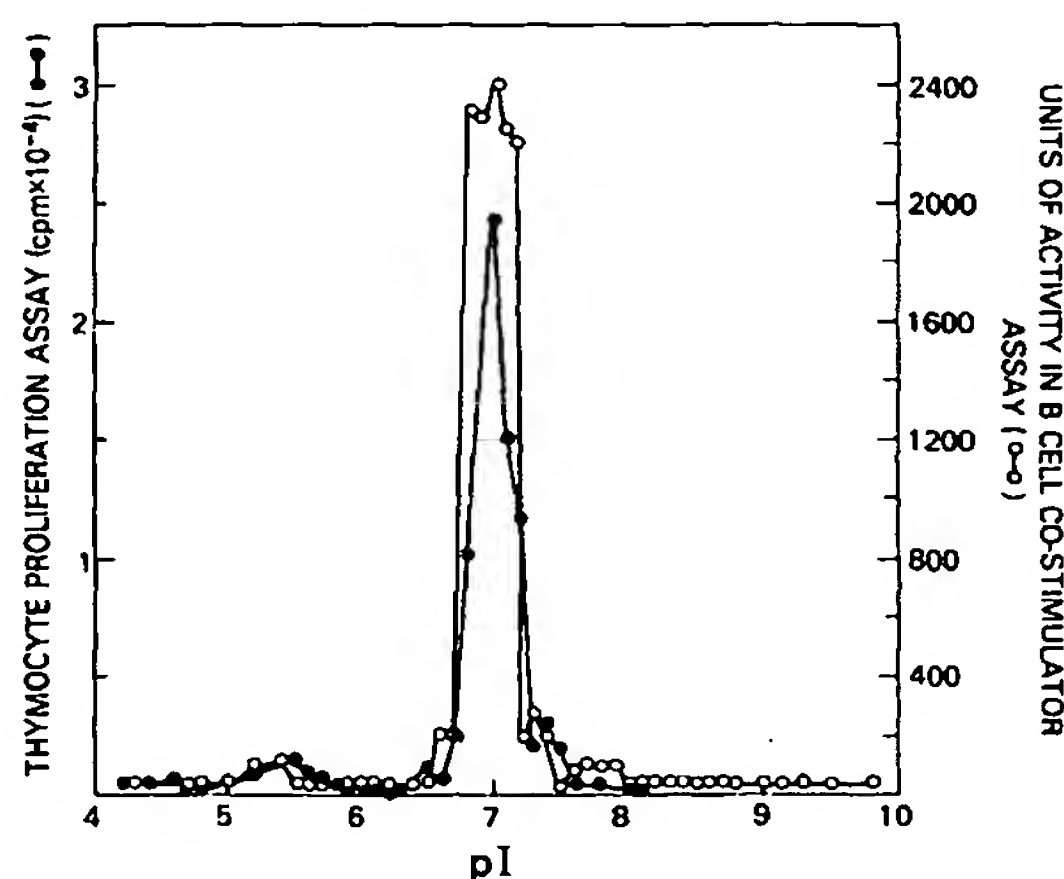


FIG. 6. B cell co-stimulating activity of partially purified human IL-1. Supernatants from endotoxin-stimulated human monocytic leukemia cells were fractionated by hollow fiber filtration, and the 10-50,000-mol wt component further purified by IEF. Dialyzed fractions were then added at several dilutions ranging from 1/50 to 1/3,200 to cultures of  $2 \times 10^4$  B cells, 10% EL4 supernatant, and anti-IgM (5  $\mu$ g/ml) (○), or at 1:100 (final dilution) to cultures of  $1.5 \times 10^6$  mouse thymocytes (●). In both cases, proliferation was assessed at day 2.5 by [ $^3$ H]thymidine uptake. Units of B cell co-stimulator activity were calculated as outlined in Fig. 5. All results represent duplicate cultures.

assayed for M15 and IL-1 activity. The two activities had a concordant distribution with a peak in activity in the pI 7.0 fractions (Fig. 6). The pI 6.8-7.2 fractions from IEF that contained both M15 and IL-1 activity were pooled and electrophoresed on



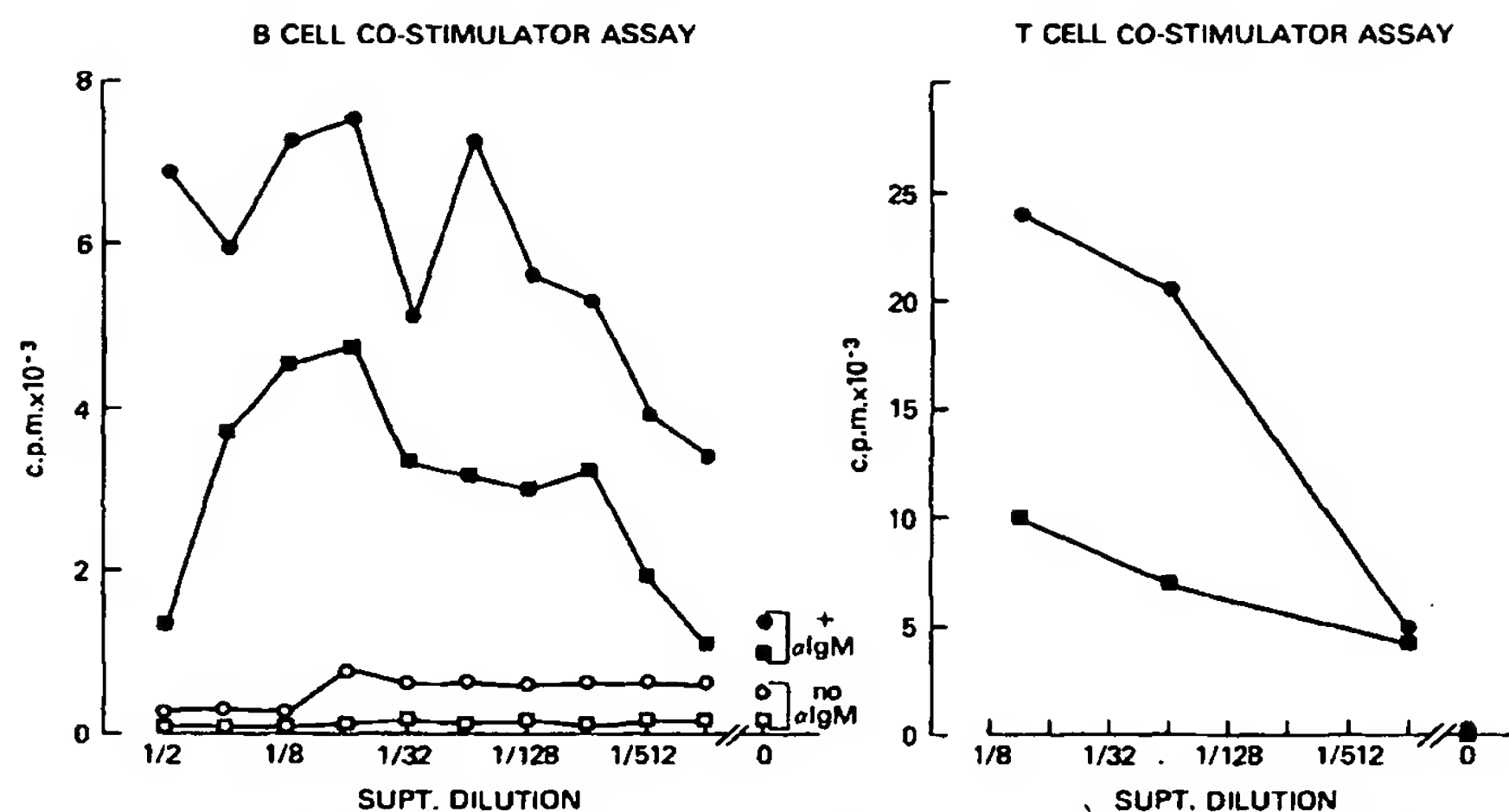


FIG. 7. B cell co-stimulating activity of purified human IL-1. Two separate preparations (●, ■) of human IL-1 from endotoxin-stimulated monocytic leukemia cells were purified to high specific activity as described in text. Each preparation was added at the various dilutions indicated to cultures of  $2 \times 10^4$  B cells, 10% EL4 supernatant with or without anti-IgM at 5  $\mu$ g/ml (B cell co-stimulator assay), or mouse thymocytes (T cell co-stimulator assay). Proliferation was assessed at day 2.5 by [ $^3$ H]thymidine uptake. Results represent means of duplicate cultures.

a tris glycinate polyacrylamide gel. In two separate experiments, the fraction that contained IL-1 activity was shown to be also rich in M15 activity (Fig. 7).

The biochemical results correlating B cell and thymocyte co-stimulating activities from two separate sources of IL-1 provide strong evidence that M15 is the monokine IL-1 and thus that IL-1, together with BCGF, plays a key role in the stimulation of anti-IgM treated B cells to enter the S phase of the cell cycle.

### Discussion

We report here conditions for polyclonal activation of small numbers of highly purified mouse B lymphocytes. Three stimuli are required for induction of DNA synthesis by the particular subset of small B lymphocytes investigated: one delivered by antibodies specific for the IgM receptor expressed on the B cell membrane; one delivered by the T cell-derived factor BCGF; and one delivered by the 15,000-mol wt macrophage-derived factor IL-1. BCGF is a newly described lymphokine discussed in detail elsewhere (9, 12).<sup>2</sup> Identification of the monokine involved in this process as IL-1 is based on correlation of B cell and T cell co-stimulating activities following an extensive series of biochemical purification procedures. Indeed, the purest preparations of murine and human IL-1 currently available show excellent B cell co-stimulating activity in our assay systems. We do not believe that the B cell co-stimulatory activity contained in IL-1-containing supernatants can be attributed to the PMA generally used for IL-1 induction for several reasons: (a) PMA-free IL-1-containing supernatants obtained from UV-irradiated P388D<sub>1</sub> cells contain the B cell-costimulant; (b) The IL-1-rich supernatant obtained by the superinduction protocol is produced by exposing P388D<sub>1</sub> cells to PMA for 5 h, extensively washing the cells, culturing them for 24 additional h and collecting that supernatant. No PMA can be detected in the purified

preparation, as assessed by the inclusion of isotopically labeled PMA in test batches of supernatant (Mizel, unpublished data); (c) The human IL-1-containing supernatant is induced by endotoxin rather than by PMA; and (d) The B cell co-stimulant resided in IL-1 purified to apparent homogeneity. We wish to emphasize that BCGF and IL-1 do not appear to be mitogenic for resting B cells. Furthermore, BCGF derived from an H-2<sup>b</sup> line (EL4) functions on BALB/c (H-2<sup>d</sup>) B cells and IL-1 derived from an H-2<sup>d</sup> line (P388D<sub>1</sub>) functions on CBA (H-2<sup>k</sup>) B cells (unpublished data). Thus, these factors act on anti-IgM-activated B cells in a non-antigen-specific, non-H-2-restricted manner. The fact that BCGF and IL-1 co-stimulatory effects are synergistic rather than additive suggests the two factors operate on the same B cell subset. We show elsewhere that BCGF does not lead to the appearance of antibody-forming cells, but is nevertheless an essential component in both an antigen-induced antibody-forming cell response (9) and in the induction of cells with cytoplasmic Ig in response to anti-IgM.<sup>3</sup> Preliminary experiments suggest a similar role for IL-1 (unpublished observations). The involvement of accessory cells in anti-IgM-induced B cell proliferation has previously been a matter of controversy; however, involvement of either macrophages or T cells has been reported by some laboratories (28-30). We suggest that failure to observe the need for one or both cell types reflects the use of assay conditions in which small numbers of one or both accessory cells are present with consequent endogenous factor production. Our own experiments show a need for exogenous BCGF and IL-1 only at low cell densities, despite the rigorous B cell purification procedure we have used. Thus it would appear that very small numbers of accessory cells can provide B cell stimulatory co-factors.

Three major models to explain antigen-specific B cell activation have previously been proposed: (a) activation occurs as a result of the binding of antigen to membrane Ig (8, 31); (b) activation occurs via an unrelated nonspecific receptor, with membrane Ig serving to specifically focus molecules capable of binding to this receptor (32); and (c) activation requires both a signal resulting from the binding of the Ig receptor and a signal delivered by an interaction at another antigen-nonspecific receptor on the B cell membrane (33). In this study, we arrive at the novel conclusion that IgM-specific induction of B cell proliferation requires three stimuli: one specific for the antigen receptor, and two antigen-nonspecific stimuli. Preliminary kinetic analyses investigating the relative roles of BCGF and IL-1 in the B cell cycle suggest the two factors function independently, with BCGF operating on G<sub>0</sub> or early G<sub>1</sub> cells and IL-1 acting at some later point of G<sub>1</sub> (34; M. Howard and E. Rabin, unpublished data). It seems likely that the time of action of the co-stimulant correlates with that time that receptors for them are expressed on the B cell surface, just as T cell sensitivity to IL-2 appears to correlate with the expression of receptors for IL-2 on T cells that have been stimulated with lectins or antigens (35-37). However, no direct evidence for receptors for BCGF or IL-1 on activated mouse B cells has yet been obtained.

This report establishes a role for IL-1 in stimulation of B cells. While others have previously proposed roles for IL-1 in B cell responses (38-44), such experiments were conducted using less pure B cell populations cultured at high cell density and thus failed to distinguish a direct action on B cells vs. secondary actions via other cell

<sup>3</sup> Nakanishi, K., M. Howard, A. Muraguchi, J. Farrar, K. Takatsu, T. Hamaoka, and W. E. Paul. 1983. Soluble factors involved in B cell differentiation: identification of two distinct T cell replacing factors (TRF). *J. Immunol.* In press.

types. A particular problem in this respect is that IL-1 appears responsible for the generation of at least some T cell-derived lymphokines, e.g., IL-2 (22-25) and colony-stimulating factor (26). With the recent identification of several distinct B cell-specific lymphokines<sup>3</sup> (9, 16, 27, 30, 45-48), one could certainly envisage IL-1-mediated factor cascades in many biological assays for B cell function. While such objections may be leveled at our experiments, we offer the following considerations: (a) We have used very highly purified B cells cultured at low cell densities, thus greatly reducing the possibility of accessory cell contamination. (b) We show elsewhere (34) that IL-1 may be added quite late in the B cell cycle, thus providing little opportunity for a factor cascade to develop. (c) Other T cell-derived B cell-specific lymphokines, e.g., BCGF, and the T cell-replacing factors from the EL4 (EL-TRF) and the B15K12 (B15-TRF) T cell lines<sup>3</sup> fail to replace the need for IL-1 (unpublished data). Nevertheless, ultimate proof of IL-1 acting directly on B cells will await reagents capable of demonstrating its specific binding to cellular receptors. The observation that IL-1 dependency of B cell proliferation is heightened by omitting 2ME from the culture medium is reminiscent of the earlier studies of Mongini et al. (28), who demonstrated an accessory cell requirement for anti-IgM induced B cell proliferation under 2ME-free conditions. 2ME has been used extensively as an additive to culture medium in a variety of tissue culture systems to either augment the magnitude or facilitate observation of the particular phenomenon under study. While 2ME has been directly implicated as a lymphocyte mitogen and polyclonal activator (49, 50), our experiments suggest an alternative mode of action, namely as a stimulant that induces IL-1 production by residual macrophages and/or B cells.

We wish to emphasize that the conclusions reached from these studies reflect the activation requirements of the subset of B cells that can proliferate in response to anti-IgM antibodies. This subset has recently been quantitated as comprising ~50% of normal splenic B cells (21). Such cells are absent from *xid* mice and appear to be members of the Lyb-5<sup>+</sup> population of normal B cells. Thus, our data do not exclude the existence of a separate subset of B cells which interact with T cells in an H-2-restricted manner, as demonstrated by others (51-53). Finally, it is difficult to relate the findings of this investigation to those recently obtained in other laboratories using elegant single-cell assays for B cell growth (54-56). As such studies have not used anti-IgM antibodies as the primary activant, different B cell subsets may be under study. Furthermore, the increased sensitivity of single-cell assays over those described in this report may allow detection of trace amounts of growth factors in cell supernatants or indeed in the fetal calf serum contained in culture medium itself, producing results in apparent conflict with ours. Attempts to resolve some of these issues are in progress.

### Summary

In this report we describe conditions for polyclonal activation of small numbers of highly purified mouse B lymphocytes. Three signals are required for induction of DNA synthesis by the particular subset of small B lymphocytes investigated: a signal delivered by antibodies specific for the IgM receptor expressed on the B cell membrane; a signal delivered by a T cell-derived factor (B cell growth factor [BCGF]); and a signal delivered by the macrophage-derived factor interleukin 1 (IL-1). The conclusion that IL-1 has B cell co-stimulator activity is based on the findings that highly purified preparations of mouse and human IL-1 have the capacity to cause

proliferation in B cells treated with anti-IgM and BCGF. Such cultures show an absolute dependence on exogenously added IL-1 when 2-mercaptoethanol is omitted from the medium. BCGF and IL-1 each act in a non-antigen-specific, non-H-2-restricted, synergistic manner. Their requirement is not observed when B cells are cultured at high density, presumably reflecting accessory cell contamination and endogenous factor production under these conditions. The B cell activation induced by these three signals is restricted to proliferation without the production of antibody-forming cells.

We thank K. Nakanishi and R. Asofsky for providing affinity-purified goat anti-mouse IgM antibodies, M. Hilfiker for assistance with gel filtration analyses, and J. Farrar for IEF-purified BCGF.

*Received for publication 18 August 1982 and in revised form 24 January 1983.*

### References

1. Sell, S., and P. G. H. Gell. 1965. Studies on rabbit lymphocytes in vitro. I. Stimulation of blast transformation with an anti-allotype serum. *J. Exp. Med.* 122:423.
2. Adenolfi, M., B. Gardner, F. Gianelli, and M. McGuire. 1967. Studies on human lymphocytes stimulated *in vitro* with anti- $\mu$  and anti- $\mu$  antibodies. *Experientia (Basel)*. 23:271.
3. Parker, D. C. 1975. Stimulation of mouse lymphocytes by insoluble anti-mouse immunoglobulins. *Nature (Lond.)*. 258:361.
4. Gausset, P., G. Delespesse, C. Hubert, B. Bennes, and A. Govaerts. 1976. *In vitro* response of subpopulations of human lymphocytes. II. DNA synthesis induced by anti-immunoglobulin antibodies. *J. Immunol.* 116:446.
5. Weiner, H. L., J. W. Moorehead, and H. Claman. 1976. Anti-immunoglobulin stimulation of murine lymphocytes. I. Age dependency of the proliferative response. *J. Immunol.* 116:1656.
6. Sieckmann, D. G., R. Asofsky, D. E. Mosier, I. M. Zitron, and W. E. Paul. 1978. Activation of mouse lymphocytes by anti-immunoglobulin. I. Parameters of the proliferative response. *J. Exp. Med.* 147:814.
7. Pure, E., and E. Vitetta. 1980. Induction of murine B cell proliferation by insolubilized anti-immunoglobulins. *J. Immunol.* 125:1240.
8. Sieckmann, D. G., I. Scher, R. Asofsky, D. E. Mosier, and W. E. Paul. 1978. Activation of mouse lymphocytes by anti-immunoglobulin. II. A thymus-independent response by a mature subset of B lymphocytes. *J. Exp. Med.* 148:1628.
9. Howard, M., J. Farrar, M. Hilfiker, B. Johnson, K. Takatsu, T. Hamaoka, and W. E. Paul. 1982. Identification of a T cell-derived B cell growth factor distinct from interleukin 2. *J. Exp. Med.* 155:914.
10. Farrar, J., J. Fuller-Farrar, P. Simon, M. Hilfiker, B. Stadler, and W. Farrar. 1980. Thymoma production of T cell growth factor (interleukin 2). *J. Immunol.* 125:2555.
11. Fuller-Farrar, J., M. Hilfiker, W. Farrar, and J. Farrar. 1981. PMA enhances the production of interleukin 2. *Cell. Immunol.* 58:156.
12. Farrar, J., W. Benjamin, M. Hilfiker, M. Howard, W. Farrar, and J. Fuller-Farrar. 1982. The biochemistry, biology, and role of IL-2 in the induction of cytotoxic T cells and antibody-forming B cell responses. *Immunol. Rev.* 63:129.
13. Mizel, S. B., D. L. Rosenstreich, and J. J. Oppenheim. 1978. Phorbol myristic acetate stimulates LAF production by the macrophage cell line, P388D<sub>1</sub>. *Cell. Immunol.* 40:230.
14. Mizel, S. B., and D. Mizel. 1981. Purification to apparent homogeneity of murine interleukin 1. *J. Immunol.* 126:834.



15. Lachman, L. B., J. O. Moore, and R. S. Metzgar. 1978. Preparation and characterization of lymphocyte-activating factor (LAF) from monocytic leukemia cells. *Cell. Immunol.* 41:199.
16. Leibson, H., P. Marrack, and J. Kappler. 1981. B cell helper factors. I. Requirement for both interleukin 2 and another 40,000-mol wt factor. *J. Exp. Med.* 154:1681.
17. Lachman, L. B., S. O. Page, and R. S. Metzgar. 1980. Purification of human lymphocyte activating factor interleukin 1. *J. Supramol. Struct.* 137:457.
18. Lachman, L. B., M. P. Hacker, and R. E. Handschumacher. 1977. Partial purification of lymphocyte-activating factor (LAF) by ultrafiltration and electrophoretic techniques. *J. Immunol.* 119:2019.
19. Gery, I., R. K. Gershon, and B. H. Waksman. 1972. Potentiation of the T-lymphocyte response to mitogens. I. The responding cell. *J. Exp. Med.* 136:128.
20. Lachman, L., M. Hacker, G. Blyden, and R. Handschumacher. 1977. Preparation of lymphocyte-activating factor from continuous murine macrophage cell lines. *Cell Immunol.* 34:416.
21. DeFranco, A., E. Raveche, R. Asofsky, and W. E. Paul. 1982. Frequency of B lymphocytes responsive to anti-immunoglobulin. *J. Exp. Med.* 155:1523.
22. Larsson, E.-L., N. N. Iscove, and A. Coutinho. 1980. Two distinct factors are required for induction of T cell growth. *Nature (Lond.)*. 283:664.
23. Smith, K. A., K. J. Gilbride, and M. F. Favata. 1980. Lymphocyte activating factors promotes T cell growth factor production by cloned murine lymphoma cells. *Nature (Lond.)*. 287:353.
24. Smith, K. A., L. B. Lackman, J. J. Oppenheim, and M. F. Favata. 1980. The functional relationship of the interleukins. *J. Exp. Med.* 151:1551.
25. Farrar, J. J., S. B. Mizel, J. Fuller-Bonar, M. L. Hilfiker, and W. L. Farrar. 1980. Lipopolysaccharide-mediated adjuvanticity: effect of lipopolysaccharide on the production of T cell growth (interleukin 2). In *Microbiology*. D. Schlessinger, editor. American Society for Microbiology, Washington, DC. 36-48.
26. Moore, R., J. Hoffeld, J. Farrar, S. Mergenhagen, J. Oppenheim, and R. Shadduck. 1981. Role of colony stimulating factors as primary regulators of macrophage functions. *Lymphokines*. 3:119.
27. Takatsu, K., K. Tanaka, A. Tominaga, Y. Kumahara, and T. Hamaoka. 1980. Antigen-induced T cell-replacing factor (TRF). III. Establishment of T cell hybrid clone continuously producing TRF and functional analysis of released TRF. *J. Immunol.* 125:2646.
28. Mongini, P., S. Friedman, and H. Wortis. 1978. Accessory cell requirement for anti-IgM proliferation of B lymphocytes. *Nature (Lond.)*. 276:709.
29. Sidman, C., and E. Unanue. 1978. Control of proliferation and differentiation in B lymphocytes by anti-Ig antibodies and a serum-derived cofactor. *Proc. Natl. Acad. Sci. USA*. 75:2401.
30. Parker, D. 1980. Induction and suppression of polyclonal antibody responses by anti-Ig reagents and antigen-nonspecific helper factors. *Immunol. Rev.* 52:115.
31. Diener, E., and M. Feldman. 1972. Relationship between antigen and antibody-induced suppression of immunity. *Transplant. Rev.* 8:76.
32. Coutinho, A., and G. Moller. 1974. Immune activation of B cells: evidence for one non-specific triggering signal not delivered by the Ig receptors. *Scand. J. Immunol.* 3:133.
33. Bretscher, P. 1975. The two signal model for B cell induction. *Transplant. Rev.* 23:37.
34. Howard, M., and W. E. Paul. 1983. Regulation of B cell growth and differentiation by soluble factors. *Ann. Rev. Immunol.* 1:307-333.
35. Lafferty, K., H. Warren, J. Woolnough, and D. Talmage. 1978. Immunological induction of T lymphocytes: role of antigen and the lymphocytes co-stimulator. *Blood Cells*. 4:395.
36. Larsson, E., and A. Coutinho. 1979. On the role of mitogenic lectins in T-cell triggering. *Nature (Lond.)*. 280:239.

37. Andersson, J., L. Gronvik, E. Larsson, and A. Coutinho. 1979. Studies on T-lymphocyte activation. I. Requirement for the mitogenic-dependent production of T-cell growth factors. *Eur. J. Immunol.* 9:581.
38. Schrader, J. W. 1973. Mechanism of activation of the bone marrow-derived lymphocyte. III. A distinction between a macrophage-produced triggering signal and the amplifying effect on triggered B lymphocytes of allogenic interactions. *J. Exp. Med.* 138:1446.
39. Wood, D. D., and S. L. Gaul. 1974. Enhancement of the humoral response of T cell-depleted murine spleens by a factor derived from human monocytes *in vitro*. *J. Immunol.* 113:925.
40. Wood, D. D., P. M. Cameron, M. T. Poe, and C. A. Morris. 1976. Resolution of a factor that enhances the antibody response of T cell-depleted murine splenocytes from several other monocyte products. *Cell. Immunol.* 21:88.
41. Wood, D. D. 1979. Purification and properties of human B cell activating factor. *J. Immunol.* 123:2395.
42. Hoffman, M. K., and J. Watson. 1979. Helper T cell-replacing factors secreted by thymus-derived cells and macrophages: cellular requirements for B cell activation and synergistic properties. *J. Immunol.* 122:1371.
43. Hoffman, M. K., S. Koenig, R. S. Mittler, H. F. Oertgen, P. Ralph, C. Galanos, and U. Hammerling. 1979. Macrophage factor controlling differentiation of B cells. *J. Immunol.* 122:497.
44. Hoffman, M. 1980. Macrophages and T cells control distinct phases of B cell differentiation in the humoral immune response *in vitro*. *J. Immunol.* 125:2076.
45. Schimpl, A., L. Hubner, C. Wong, and E. Wecker. 1980. Distinction between T helper cell replacing factor (TRF) and T cell growth factor (TCGF). *Behring Inst. Mitt.* 67:221.
46. Andersson, J., and F. Melchers. 1981. T cell-dependent activation of resting B cells: requirement for both nonspecific unrestricted and antigen-specific Ia-restricted soluble factors. *Proc. Natl. Acad. Sci. USA.* 78:2497.
47. Swain, S., G. Dennert, J. Warner, and R. Dutton. 1981. Culture supernatants of a stimulated T cell line have helper activity that synergizes with IL-2 in the response of B cells to antigen. *Proc. Natl. Acad. Sci. USA.* 78:2517.
48. Isakson, P., E. Pure, E. Vitetta, and P. Krammer. 1982. T cell-derived B cell differentiation factor(s). Effect on the isotype switch of murine B cells. *J. Exp. Med.* 155:734.
49. Lemke, H., and H. Opitz. 1976. Function of 2-mercaptoethanol as a macrophage substitute in the primary immune response *in vitro*. *J. Immunol.* 117:388.
50. Goodman, M., and W. Weigle. 1977. Nonspecific activation of murine lymphocytes. I. Proliferation and polyclonal activation induced by 2-mercaptoethanol and  $\alpha$ -thioglycerol. *J. Exp. Med.* 145:473.
51. Swierkosz, J., K. Rock, P. Marrack, and J. Kappler. 1978. The role of H-2-linked genes in helper T-cell function. II. Isolation on antigen-pulsed macrophages of two separate populations of F<sub>1</sub> helper T cells each specific for antigen and one set of parental H-2 products. *J. Exp. Med.* 147:554.
52. Sprent, J. 1978. Restricted helper function of F<sub>1</sub> hybrid T cells positively selected to heterologous erythrocytes in irradiated parental strain mice. I. Failure to collaborate with B cells of the opposite parental strain not associated with active suppression. *J. Exp. Med.* 147:1142.
53. Singer, A., P. Morrissey, H. Hathcock, A. Ahmed, I. Scher, and R. Hodes. 1981. Role of the major histocompatibility complex in T cell subpopulations differ in their requirement for major histocompatibility complex-restricted T cell recognition. *J. Exp. Med.* 154:501.
54. Wetzel, G. D., and J. R. Kettman. 1981. Activation of murine B lymphocytes. III. Stimulation of B lymphocyte clonal growth with lipopolysaccharide and dextran sulfate. *J. Immunol.* 126:723.

55. Vaux, D. L., B. L. Pike, and G. J. V. Nossal. 1981. Antibody production by single hapten-specific B lymphocytes: an antigen-driven cloning system free of filler or accessory cells. *Proc. Natl. Acad. Sci. USA.* 78:7702.
56. Wetzel, G., S. Swain, and R. Dutton. 1982. A monoclonal T cell replacing activity can act directly on B cells to enhance clonal expansion. *J. Exp. Med.* 156:306.

# BINDING AND INTERNALIZATION OF INTERLEUKIN 1 BY T CELLS

## Direct Evidence for High- and Low-Affinity Classes of Interleukin 1 Receptor

BY JOHN W. LOWENTHAL AND H. ROBSON MACDONALD

*From the Ludwig Institute for Cancer Research, Lausanne Branch, 1066 Epalinges,  
Switzerland*

IL-1, a macrophage-derived 17,500  $M_r$  glycoprotein, has been shown to mediate a wide spectrum of biological activities (1, 2). Insofar as T cell-mediated immunity is concerned, IL-1 has been shown in several studies (3-9) to induce Th cell secretion of the growth hormone IL-2, while in other systems it has been shown to provide an obligatory signal for IL-2-R expression (10, 11). The study of the regulatory role of IL-1 in T cell proliferation has been greatly facilitated by the recent development of a direct radiolabeled IL-1 binding assay by Dower et al. (12). These authors showed that IL-1 bound to specific cell surface receptors that have a molecular weight of around 80,000. The general level of IL-1-R expression by a variety of cell types was found to be much lower than that of receptors for other hormones, such as epidermal growth factor, insulin, and IL-2. Various T cell populations, for example, were shown to express only 27-550 IL-1-R per cell (12).

We have recently characterized a mutant subline of the EL4 thymoma (EL4-6.1), which is induced to simultaneously secrete IL-2 and express IL-2-R in the presence of IL-1 (13). We show in the present report, using a direct IL-1 binding assay that these cells express extremely high numbers of IL-1-R (~20,000 per cell). The IL-1-R expressed by EL4-6.1 cells can be resolved into two classes: the vast majority of receptors bind IL-1 with a  $K_d$  of ~200-500 pM, whereas a second class of receptor, making up 1-2% of the total, binds IL-1 with a 100-fold higher affinity ( $K_d$  of 3-8 pM). This latter value is equivalent to the concentration of IL-1 that gives 50% biological activity. In addition, the two classes of IL-1-R can be distinguished on the basis of their ability to internalize IL-1; it appears that only high-affinity IL-1-R can do so. Other cell types also express both classes of IL-1-R, but the absolute number of receptors per cell is considerably less than on EL4-6.1 cells. Because of their high degree of responsiveness to IL-1 and the expression of unusually high numbers of IL-1-R, EL4-6.1 cells offer a valuable system with which regulation of IL-1-R expression and the mechanism of IL-1 action can be studied in detail.



### Materials and Methods

**Cell Cultures.** The murine T cell lines used in this study were: EL4 thymoma sublines EL4-6.1 (14), EL4-10 (subclone of EL4-6.1), EL4-3 (an independent non-IL-2-secreting subline), EL4-RN, and EL4-RP (kindly provided by O. Kanagawa, Lilly Research Laboratories, La Jolla, CA); the IL-1-responsive T lymphoma LBRM-33-1A5B6 (LBRM, reference 15); the IL-2-dependent cytolytic T cell line CTLL (16); and the T lymphomas ST-4.2 (17), BW5147 (18), and Yac (19). Nylon wool-purified peripheral T cells from lymph nodes (LNT)<sup>1</sup> were prepared as described (20) and were >95% Thy-1<sup>+</sup> as judged by FACS analysis. Cells were cultured in enriched DME (21) and maintained at 37°C in a humidified atmosphere of 5% CO<sub>2</sub> in air. PMA was purchased from Sigma Chemical Company, St. Louis, MO. Ionomycin was obtained from Calbiochem-Behring Corp., La Jolla, CA. Human rIL-2 was provided by Biogen SA, Geneva, Switzerland, and was radioiodinated according to the method of Robb et al. (22).

**Interleukin 1.** Human recombinant IL-1 ( $\alpha$  and  $\beta$  forms, reference 23) was kindly provided by Dr. C. Henney, Immunex Corp., Seattle, WA. A detailed description of their preparation and purity is reported elsewhere (24). The specific activity of rIL-1- $\alpha$  and - $\beta$  was 10<sup>7</sup> U/mg in both the murine thymocyte assay (1, 3) and in the EL4-6.1 IL-2 induction assay (13).

**Biological Assay for IL-1.** The biological activity of IL-1 was measured by its ability to stimulate IL-2 production by EL4-6.1 cells in the presence of Ionomycin (13). EL4-6.1 cells (10<sup>5</sup> cells/microwell) were cultured in the presence of various concentrations of IL-1, and 0.1  $\mu$ g/ml of Ionomycin. After 24 h, the cell-free supernatant was measured for IL-2 activity using the CTLL line (16), according to the method of Landegren (25). In this assay, 1 U/ml of IL-2 supports 50% maximal proliferation during a 48-h culture period. In some experiments, IL-1 biological activity was measured by its ability to induce IL-2-R expression by EL4-6.1 cells in the presence of suboptimal concentrations of PMA. Briefly, EL4-6.1 cells (5  $\times$  10<sup>5</sup> cells/ml) were cultured in the presence of 0.3 ng/ml of PMA plus various concentrations of IL-1. After 48 h, expression of IL-2-R was measured by using a direct IL-2 binding assay as described previously (26, 27).

**IL-1 Absorption Assay.** Various cell lines were extensively washed and then incubated (10<sup>7</sup> cells/ml) at 4°C in the presence of a known concentration of rIL-1. After 5 h, the cell-free supernatant was measured for residual IL-1 activity in both the IL-2 production and IL-2-R expression assays using EL4-6.1 cells (13).

**Radioiodination of rIL-1.** rIL-1- $\alpha$  was radioiodinated using the chloramine T method (28). The method used was similar to that described for IL-2 by Robb et al. (22), and was according to the procedure of Dr. S. K. Dower, with some modifications. Small quantities of rIL-1- $\alpha$  (500 ng in 50  $\mu$ l of PBS pH 7.2) were added to a mixture of 15  $\mu$ l of chloramine T (30  $\mu$ g/ml in PBS) and 1 mCi (10  $\mu$ l) of Na<sup>125</sup>I (Amersham International, Amersham, United Kingdom). The reaction was carried out in an Eppendorf tube at 4°C for 10 min. The labeled rIL-1 (<sup>125</sup>I-IL-1) was separated from free Na<sup>125</sup>I on a Sephadex G-25 column (PD-10; Pharmacia Fine Chemicals, Uppsala, Sweden). ~90% of the radioactivity present in the pooled fractions 5-11 (Fig. 1), was TCA-precipitable. Biological activity of <sup>125</sup>I-IL-1 was measured as described above. ~60-70% of the initial protein was recovered as measured from experiments using mock-labeled IL-1. The specific activity of the <sup>125</sup>I-IL-1 was 10<sup>6</sup> cpm/ng and produced a single band of 17,000 M<sub>r</sub> when analyzed by SDS-PAGE.

**<sup>125</sup>I-IL-1 Binding Assay.** Binding of <sup>125</sup>I-IL-1 was measured according to the method of Dower et al. (12) and our previously published procedures for radiolabeled IL-2 (27). Briefly, aliquots of 10<sup>6</sup> cells were incubated in the presence of various concentrations of <sup>125</sup>I-IL-1 at 4°C for various periods of time (equilibrium binding occurs within 2-4 h). Free radioactivity was separated from bound by centrifugation through an oil gradient (26). Nonspecific binding was measured in the presence of a 50- to 100-fold excess of unlabeled IL-1. Scatchard plot analysis of equilibrium binding data was performed as previously described (27).

**Internalization of IL-1.** Cells were washed thoroughly and incubated (5  $\times$  10<sup>6</sup> cells/ml)

<sup>1</sup> Abbreviation used in this paper: LNT, nylon wool-purified lymph node T cells.

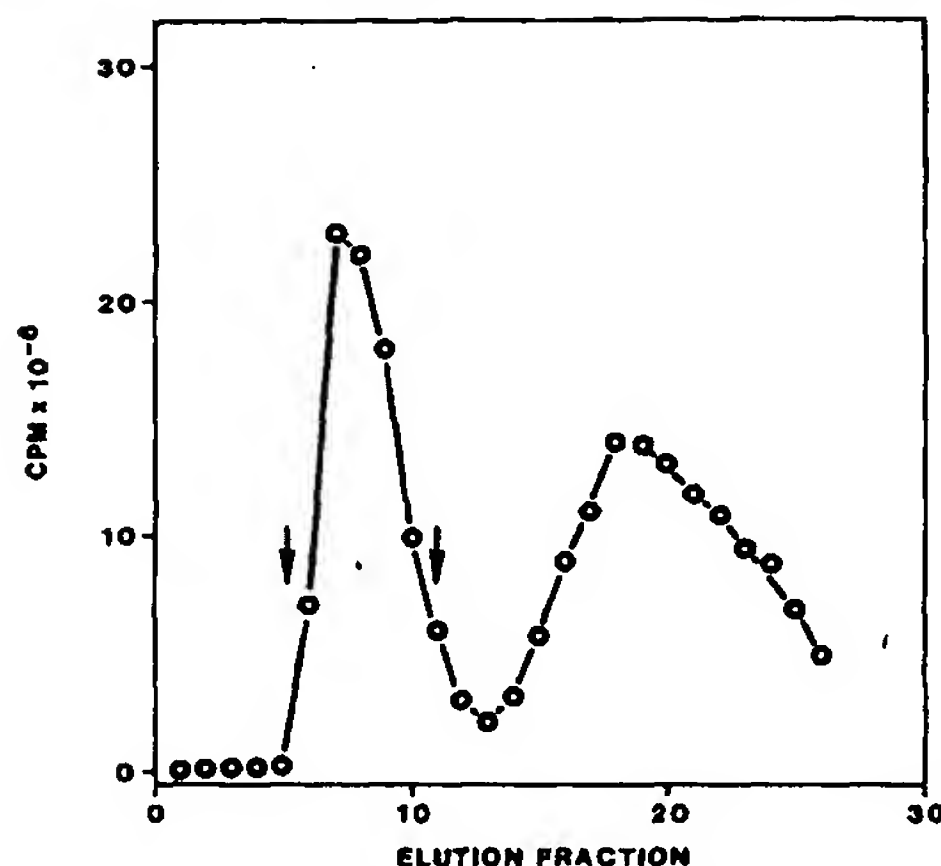


FIGURE 1. Separation of  $^{125}\text{I}$ -IL-1 from free  $\text{Na}^{125}\text{I}$ . The  $^{125}\text{I}$ -IL-1-reaction mixture was placed on top of a 1-ml Sephadex G-25 column that had been pretreated with BSA (10 mg/ml) and then extensively washed with PBS.  $^{125}\text{I}$ -IL-1 was eluted with PBS (pH 7.2) and 50  $\mu\text{l}$  fractions were collected and measured for radioactivity. Elution fractions 5-11 (arrows) were pooled and diluted in medium containing 10% FCS as carrier protein.

in the presence of  $^{125}\text{I}$ -IL-1 at  $4^\circ\text{C}$  for 3-4 h. The cells were then either washed extensively in medium or not washed and then transferred to  $37^\circ\text{C}$ . The amount of internalized IL-1 and total cell-associated IL-1 was measured at various times by centrifugation of cells through an oil gradient. Internalized IL-1 was distinguished from surface-bound IL-1 by its resistance to treatment with tissue culture medium adjusted to pH 3 by adding 1N HCl. At pH 3, >95% of surface-bound IL-1 dissociated within 1 min, whereas the level of intracellular IL-1 was not affected. The rate of increase in the pH 3-resistant IL-1 binding with time at  $37^\circ\text{C}$  was therefore taken as a measure of the rate of IL-1 internalization. The level of radioactivity remaining after pH 3 treatment at the end of the  $4^\circ\text{C}$  incubation period was taken as background. No IL-1 internalization occurred at  $4^\circ\text{C}$ . Internalization of radiolabeled rIL-2 was measured according to the same method. In other experiments, the exclusively internal localization of pH 3-resistant radioactive ligand was directly confirmed by quantitative autoradiography on electron microscopy sections (J. Lowenthal and B. Iacopetta, unpublished observations).

## Results

**Identification of IL-1- $R^+$  Cells by Absorption Experiments.** In preliminary studies we used IL-1 absorption experiments to distinguish between IL-1- $R^+$  and IL-1- $R^-$  cell types. In these experiments, cells were incubated at  $4^\circ\text{C}$  in the presence of a known concentration of rIL-1 for 4-6 h. The cell-free supernatants were then measured for residual IL-1 activity, either by their ability to induce IL-2 secretion or IL-2-R expression by EL4-6.1 cells. Fig. 2 shows that cells from IL-1-responsive lines could absorb IL-1 activity, whereas nonresponsive cell lines such as CTLL and ST4 could not. This was not due to secretion of inhibitory products by these cells, as revealed by mixing experiments. EL4-6.1 and cells of its subclone EL4-10 were very efficient at absorbing IL-1 activity. Results from experiments in which increasing numbers of EL4-6.1 cells were incubated in the presence of a constant concentration of IL-1 revealed that these cells bound about 10,000-20,000 IL-1 molecules per cell (data not shown).

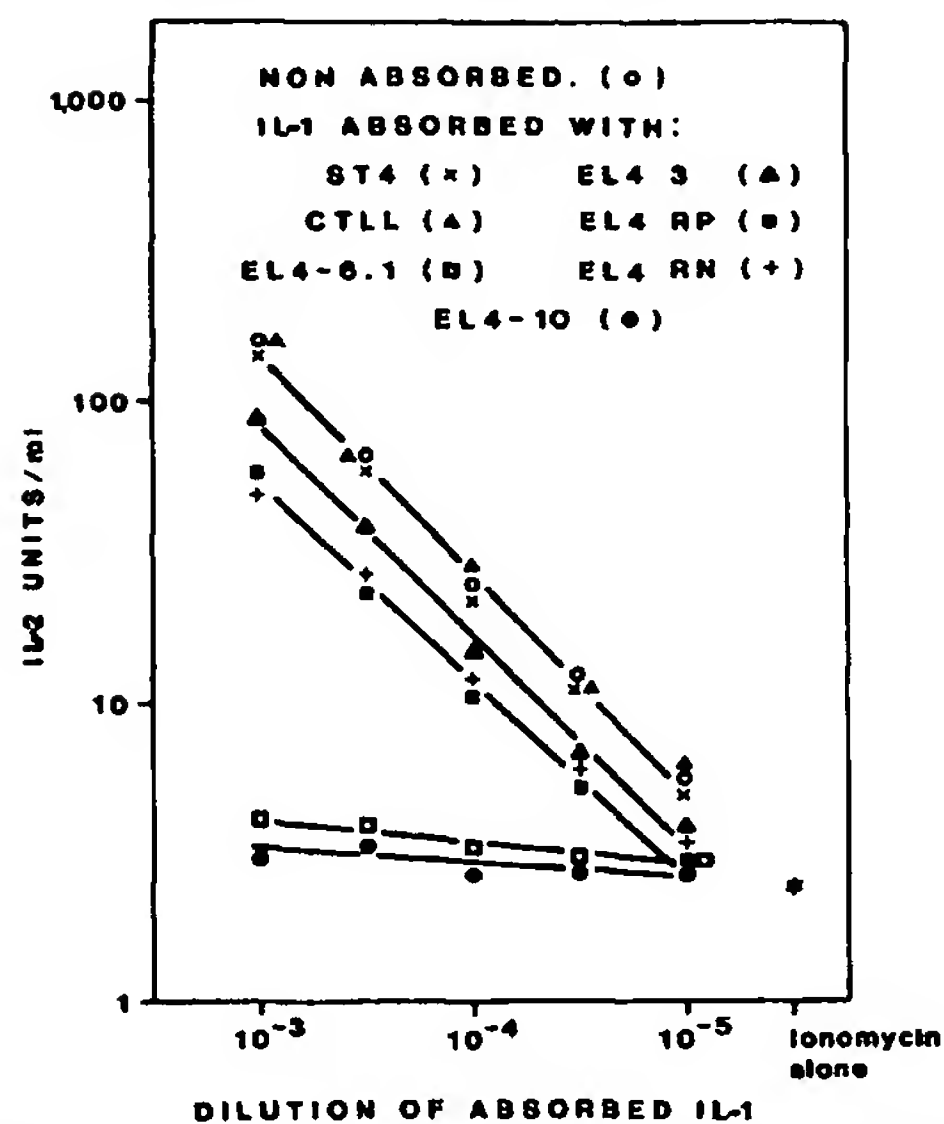


FIGURE 2. Absorption of rIL-1 by various cell lines.  $10^7$  cells/ml were incubated in the presence of rIL-1- $\alpha$  (5 ng/ml) for 4 h at 4°C. The IL-1 content of the cell-free supernatant was then measured at the indicated dilutions by its ability to stimulate IL-2 production by EL4-6.1 cells in the presence of 0.1  $\mu$ g/ml Ionomycin. Similar results were obtained using rIL-1- $\beta$  or a mixture of  $\alpha$  and  $\beta$ .

**Radioiodinated rIL-1 Retains Biological Activity.** IL-1- $\alpha$  retained full biological activity after radioiodination, as measured by its ability to induce IL-2 secretion by EL4-6.1 cells in the presence of Ionomycin (Fig. 3A). We saw half-maximum biological activity at an IL-1 concentration of 6 pM. Furthermore, Fig. 3B shows that both  $\alpha$  and  $\beta$  forms of unlabeled IL-1 competed with  $^{125}$ I-IL-1- $\alpha$  for binding, although IL-1- $\beta$  had about threefold lower affinity. Competition for  $^{125}$ I-IL-1- $\alpha$  binding by IL-1 was specific since <5% inhibition of binding was observed when unlabeled rIL-2 was added at concentrations of up to 100 ng/ml (i.e., a 1,000-fold molar excess; Fig. 3B).

**Association Kinetics of  $^{125}$ I-IL-1 with EL4-6.1 Cells.** Fig. 4A shows the time course of association of  $^{125}$ I-IL-1 with EL4-6.1 cells at 4°C over a 100-fold range in IL-1 concentrations. Binding reached equilibrium within 2–4 h over the range of concentrations tested. Logarithmic conversion of these data revealed a first order association rate constant, which was similar at all three concentrations, with a  $t_{1/2}$  value of 30–40 min. Fig. 4B shows that binding of  $^{125}$ I-IL-1 to EL4-6.1 cells occurred more rapidly at 37°C than at 4°C. The difference in association rate constant was about fourfold. Dissociation of IL-1 occurred slowly at 4°C ( $t_{1/2} > 4$  h), which is in agreement with the results of Dower et al. (12).

**Equilibrium Binding Analysis Reveals Two Classes of IL-1-R.** EL4-6.1 cells were incubated at 4°C in the presence of various concentrations of  $^{125}$ I-IL-1. After equilibrium binding was established (4 h), the cell-bound radioactivity was separated from the free radioactivity by centrifugation through an oil gradient, and

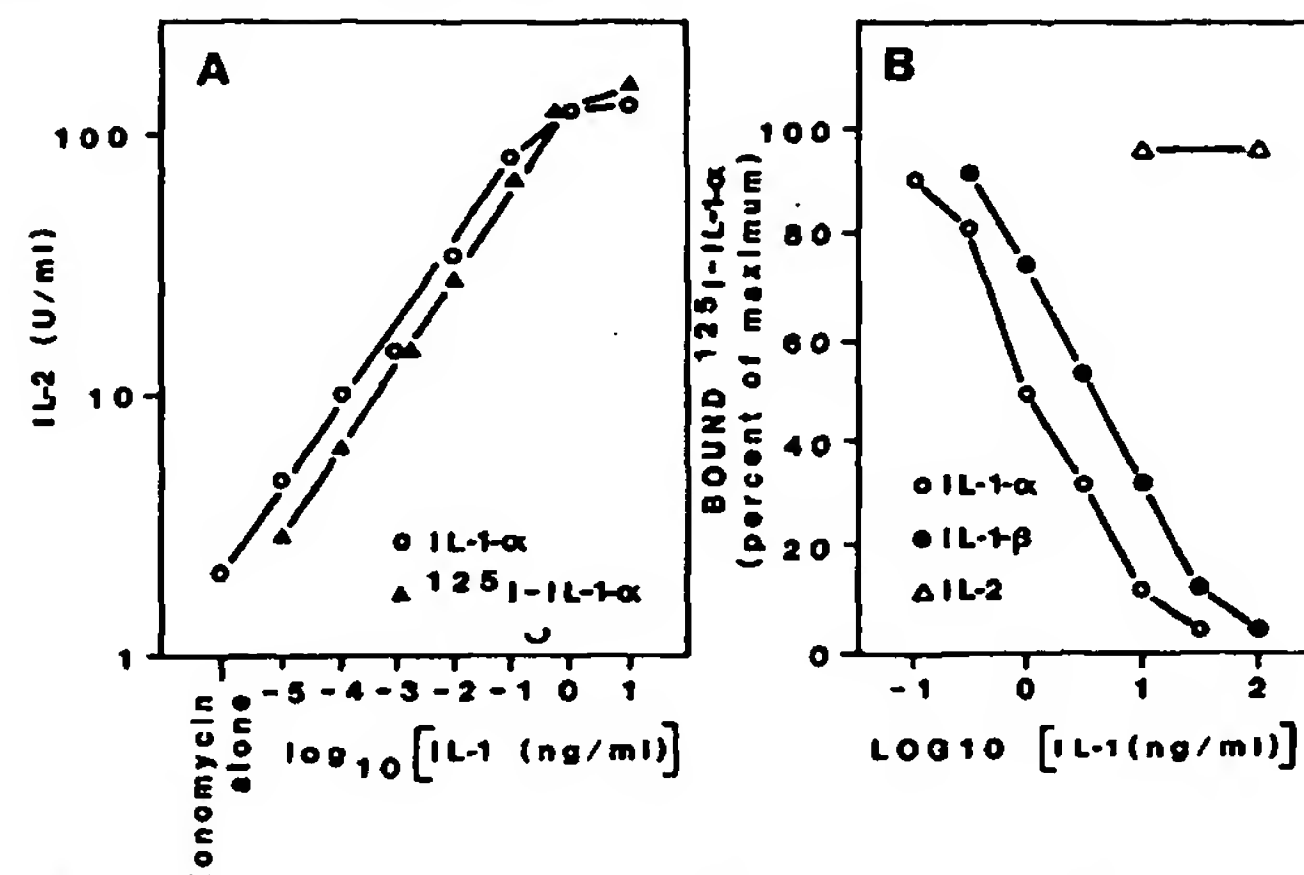


FIGURE 3.  $^{125}$ I-IL-1 retains biological activity. (A) Comparison of the capacity of labeled and unlabeled rIL-1- $\alpha$  to induce IL-2 secretion by EL4-6.1 cells in the presence of Ionomycin. (B) Ability of  $^{125}$ I-IL-1- $\alpha$  to compete for binding with unlabeled IL-1- $\alpha$  or IL-1- $\beta$ . EL4-6.1 cells ( $10^6$  cells/well) were incubated in the presence of  $^{125}$ I-IL-1- $\alpha$  (0.1 ng/ml) and various concentrations of unlabeled IL-1- $\alpha$  or IL-1- $\beta$  for 4 h at 4°C. Bound radioactivity was measured as described in Materials and Methods. Maximum radioactivity was 3,660 cpm. Unlabeled rIL-2 was also added as a negative control.

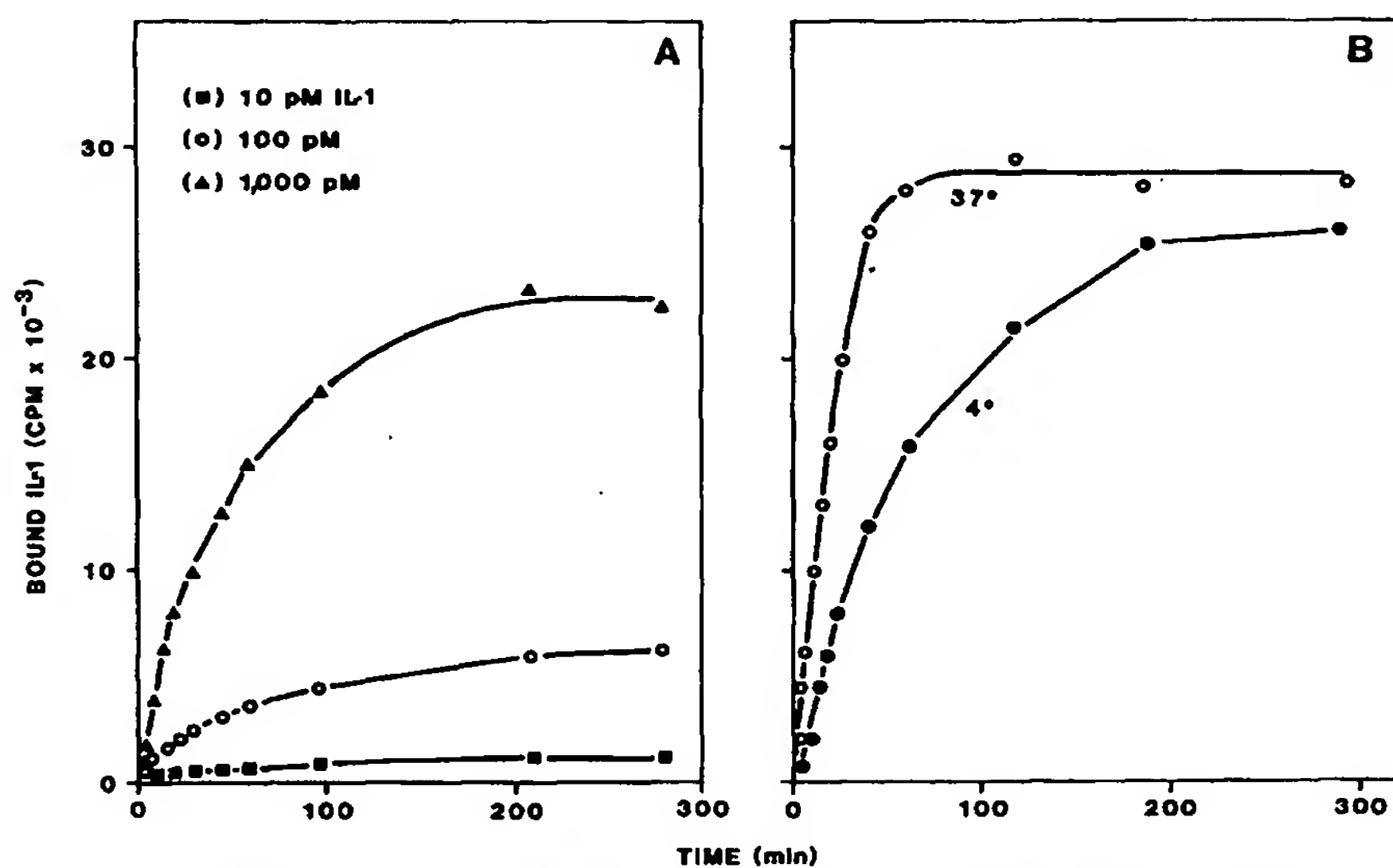


FIGURE 4. Association kinetics of  $^{125}$ I-IL-1 with EL4-6.1. (A) Cells were incubated at 4°C in the presence of the indicated concentrations of  $^{125}$ I-IL-1 for various periods of time. (B) Cells were incubated at 4°C or 37°C in the presence of  $^{125}$ I-IL-1 (1 nM) and sodium azide (0.02%) for various periods of time. Nonspecific binding has been subtracted in A and B.



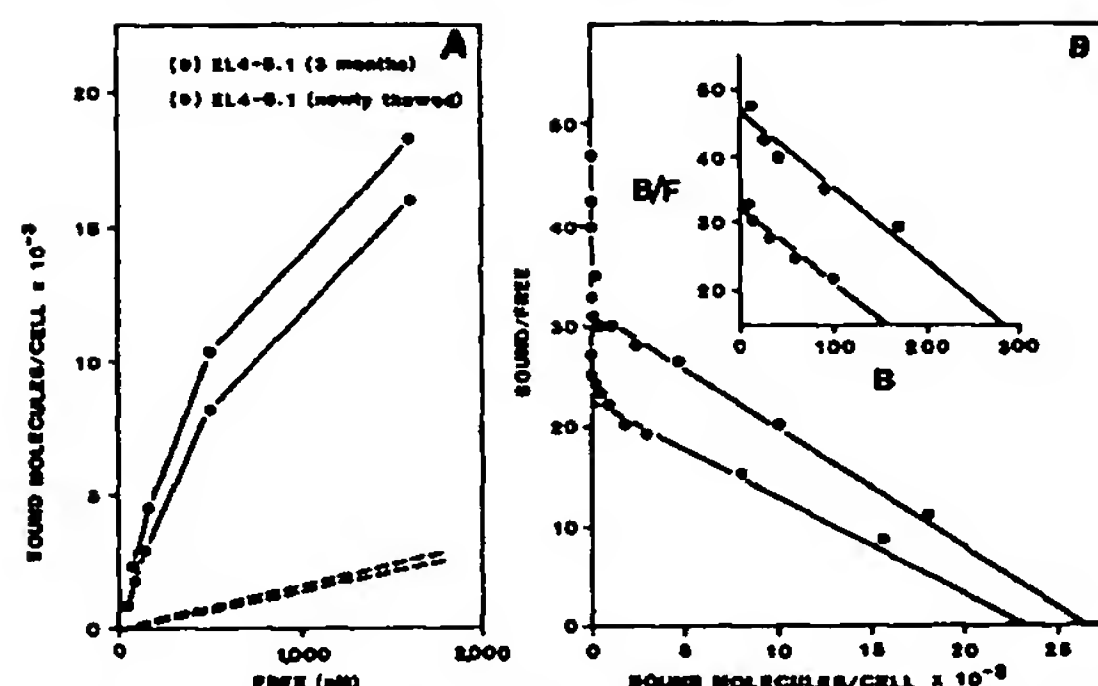


FIGURE 5. Equilibrium binding analysis of  $^{125}\text{I}$ -IL-1 to EL4-6.1. Cells were incubated at  $4^\circ\text{C}$  in the presence of various concentrations of  $^{125}\text{I}$ -IL-1 (0.2–2,000 pM) for 4 h. Freshly thawed EL4-6.1 cells (○) were compared to EL4-6.1 cells that had been cultured for 3 months (●) in their ability to bind  $^{125}\text{I}$ -IL-1. (A) Specific equilibrium binding of  $^{125}\text{I}$ -IL-1 (solid lines) after subtraction of nonspecific binding (dashed lines). (B) Scatchard plot analysis of the binding data shown in A. Inset shows the binding of  $^{125}\text{I}$ -IL-1 to the high-affinity class of IL-1-R in more detail. Similar results were obtained in six independent experiments.

TABLE I  
Expression of High- and Low-Affinity IL-1-R by Different Types of T Cells

Cells	Number of experiments	IL-1-R per cell (mean $\pm$ SEM)		Dissociation constant (mean $\pm$ SEM; pM)	
		High affinity	Low affinity	High affinity	Low affinity
EL4-6.1	5	340 $\pm$ 120	17,700 $\pm$ 4,890	4.6 $\pm$ 2.5	380 $\pm$ 140
EL4-16	2	280 $\pm$ 80	11,100 $\pm$ 3,330	3.9 $\pm$ 1.8	440 $\pm$ 170
EL4-RP	2	195 $\pm$ 105	5,700 $\pm$ 1,060	4.7 $\pm$ 1.2	335 $\pm$ 110
EL4-3	5	<2	<2		
LBRM	4	46 $\pm$ 18	2,550 $\pm$ 660	5.3 $\pm$ 2.7	360 $\pm$ 160
Yac	2	35 $\pm$ 16	1,650 $\pm$ 300	3.3 $\pm$ 1.2	290 $\pm$ 60
CTLL	5	<2	<2		
BW5147	3	<2	<2		
ST4	3	<2	<2		
LNT*	2	12 $\pm$ 4	180 $\pm$ 40	5.0 $\pm$ 1.5	405 $\pm$ 145
LNT day 2†	2	23 $\pm$ 6	580 $\pm$ 180	5.5 $\pm$ 1.8	340 $\pm$ 90

\* Freshly isolated LNT.

† LNT cultured 2 d in the presence of PMA (1 ng/ml) and Ionomycin (0.25  $\mu\text{g}/\text{ml}$ ).

then quantitated. Fig. 5A shows the specific equilibrium binding after subtraction of the nonspecific binding. Scatchard plot analysis of these data (Fig. 5B) revealed that EL4-6.1 cells express two classes of IL-1-R. ~98–99% of the total IL-1-R (~18,000/cell) bound IL-1 with a  $K_d$  of  $380 \pm 140$  pM (Table I). A second class of IL-1-R that is expressed at a low level ( $340 \pm 120$ /cell) bound IL-1 with a much higher affinity ( $K_d = 4.6 \pm 2.5$  pM; Table I and Fig. 5B, inset). IL-1-R expression by EL4-6.1 cells was stable over a 3-mo culture period (Fig. 5).

The level of high-affinity IL-1-R expressed by EL4-6.1 cells was compared with that of other T cell lines, as well as with normal peripheral T cells (Fig. 6 and Table I). Other, independently derived EL4 cell lines varied in their expression of IL-1-R, which correlated well with their ability to respond to IL-1

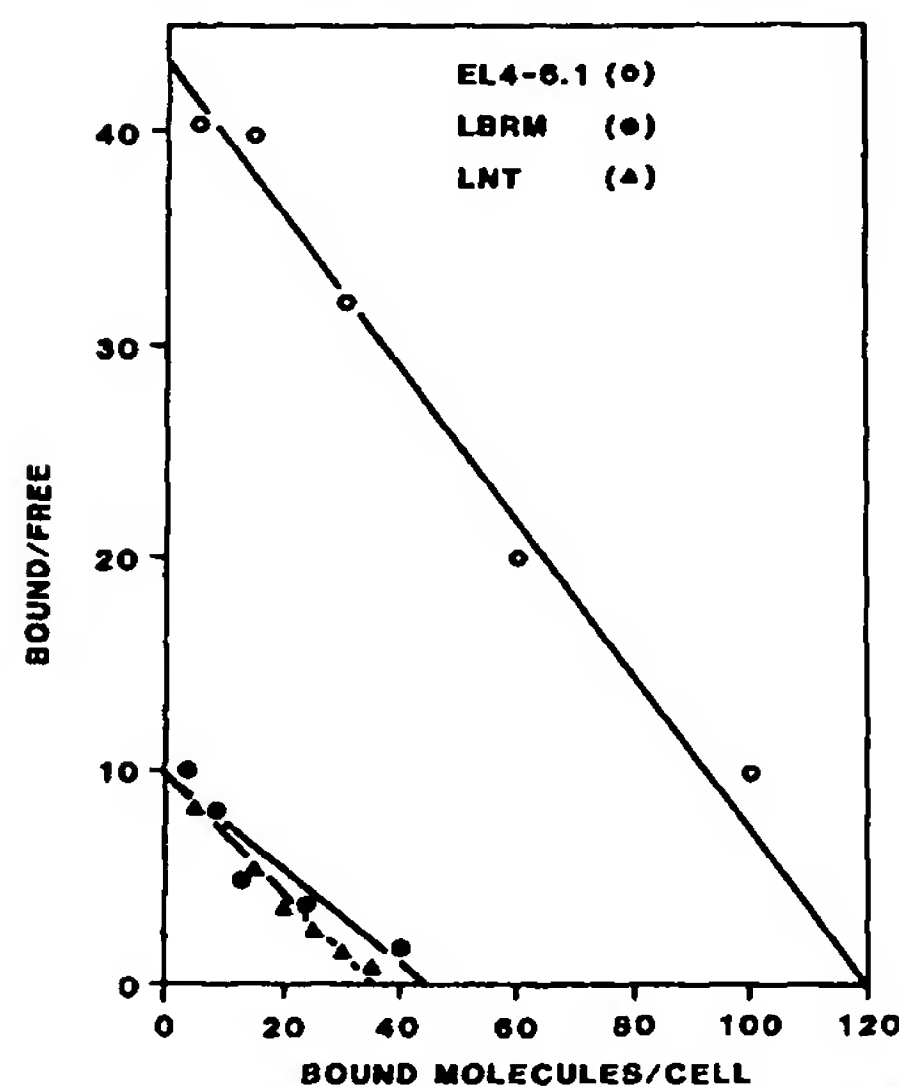


FIGURE 6. Quantitation of high-affinity IL-1-R expressed by different cell types. EL4-6.1 (○), LBRM (●) and LNT cultured for 2 d in the presence of PMA and Ionophore, (▲) were incubated in the presence of various concentrations of  $^{125}\text{I}$ -IL-1 (0.2–10 pM) for 4 h at 4°C. After subtraction of nonspecific binding, the number of binding sites per cell and the dissociation constants were quantitated by Scatchard plot analysis.

and absorb IL-1 activity. Cells of the EL4-3 line did not express IL-1-R (level of detectability is two receptors per cell), whereas EL4-RP.1 (subclone of EL4-RP) and EL4-16 (parental line of EL4-6.1) expressed intermediate levels of IL-1-R. Two other thymomas, LBRM and Yac, expressed lower numbers of IL-1-R (30–50 high-affinity IL-1-R per cell). Cells of the IL-2-dependent cytolytic T cell line CTLL, and two other thymomas, ST4 and BW5147, were negative for IL-1-R expression as was already suggested by their inability to absorb IL-1 activity (Fig. 2). In addition, activated normal T cells expressed a total of ~600 IL-1-R per cell, of which about 20 were of the high-affinity type. Interestingly, normal resting T cells also expressed detectable numbers of IL-1-R (~10 high-affinity and 180 low-affinity receptors per cell).

*Internalization of  $^{125}\text{I}$ -IL-1.* We used the ability to internalize IL-1 as an independent criterion for the expression of functional IL-1-R. EL4-6.1 cells were incubated at 4°C in the presence of  $^{125}\text{I}$ -IL-1 for 4 h to achieve equilibrium binding of IL-1 to both high- and low-affinity IL-1-R. Treatment with pH 3-buffered medium resulted in the rapid dissociation (95–98% within 1 min) of surface-bound IL1. The remaining radioactivity was taken as background. When the cells were transferred to 37°C there was a time- and temperature-dependent increase in the proportion of cell-bound radioactivity that was resistant to acid treatment. This represents IL-1 that had been internalized. Internalization is a specific process, which requires the expression of IL-1-R. Fig. 7 shows that EL4-6.1 cells (IL2-R<sup>-</sup>) had the capacity to internalize  $^{125}\text{I}$ -IL-1, but failed to inter-

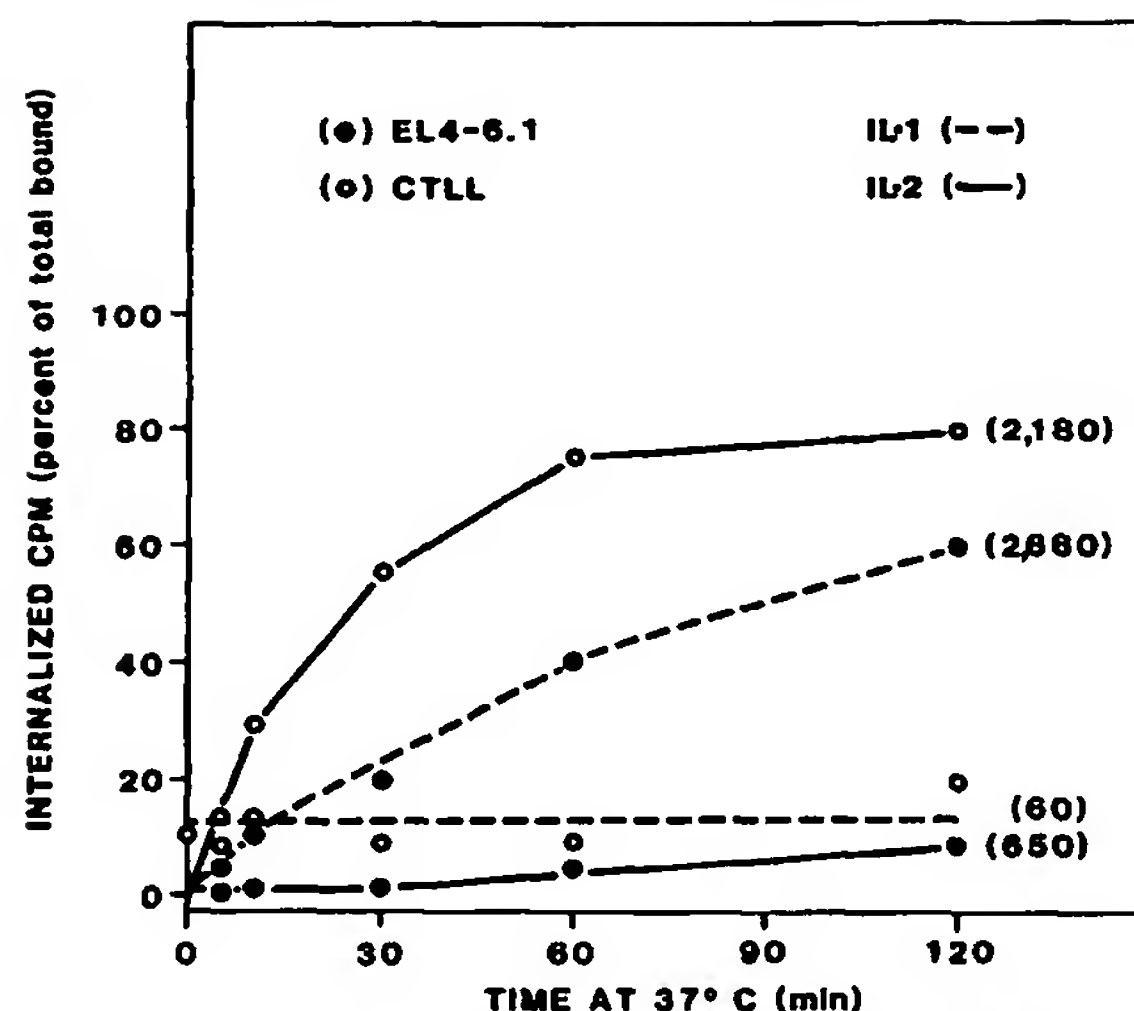


FIGURE 7. Capacity to internalize IL-1 correlates with the expression of specific IL-1-R. EL4-6.1 (●) and CTLL (○) cells were incubated at 4°C in the presence of 30 pM of  $^{125}\text{I}$ -IL-1 (dashed lines) or  $^{125}\text{I}$ -IL-2 (solid lines). After 4 h the cells were transferred to 37°C and internalized radioactivity was measured at different times, as described in Materials and Methods. Internalized IL-1 and IL-2 is expressed as a percentage of the total cell-associated ligand. The numbers in brackets refer to the total cell-bound cpm at time 0.

nalize  $^{125}\text{I}$ -IL-2. In contrast, CTLL cells (IL-2-R<sup>+</sup>/IL-1-R<sup>-</sup>) rapidly internalized  $^{125}\text{I}$ -IL-2, but could not internalize  $^{125}\text{I}$ -IL-1.

The data shown in Fig. 8 support the hypothesis that IL-1 internalization occurs via high-affinity but not low-affinity IL-1-R. EL4-6.1 cells were incubated at 4°C in the presence of 100 pM  $^{125}\text{I}$ -IL-1 for 4 h. According to the Scatchard data shown in Fig. 5, this allowed occupancy of virtually all high-affinity receptors (~250 per cell) and ~2,000 low-affinity receptors. The cells were then washed to remove all of the unbound  $^{125}\text{I}$ -IL-1. There was no change in the level of cell-bound  $^{125}\text{I}$ -IL-1 when the cells were subsequently transferred to 37°C and incubated for an additional 4 h (the level of cell-bound  $^{125}\text{I}$ -IL-1 stayed constant because there was no rebinding of IL-1, nor was there any appreciable dissociation from either class of IL-1-R). Under these conditions, there was a rapid internalization of  $^{125}\text{I}$ -IL-1, which reached a maximum level within 20 min (Fig. 8). This represented the internalization of ~200 IL-1 molecules per cell, close to the value of initially occupied high-affinity IL-1-R. There was no additional internalization of IL-1 after this time, despite the fact that ~2,000 IL-1-R per cell were still occupied. If, on the other hand, the cells were not washed before their transfer to 37°C, the level of internalized IL-1 continued to increase with time (as does the total cell-associated IL-1). This may reflect the rebinding of free  $^{125}\text{I}$ -IL-1 to recycled or newly expressed high-affinity IL-1-R. After 60 min at 37°C, the level of internal IL-1 in the unwashed cells was three times that of the washed cells. Assuming that only the 200–300 high-affinity IL-1-R expressed by these cells can be internalized, and that these receptors return to the cell

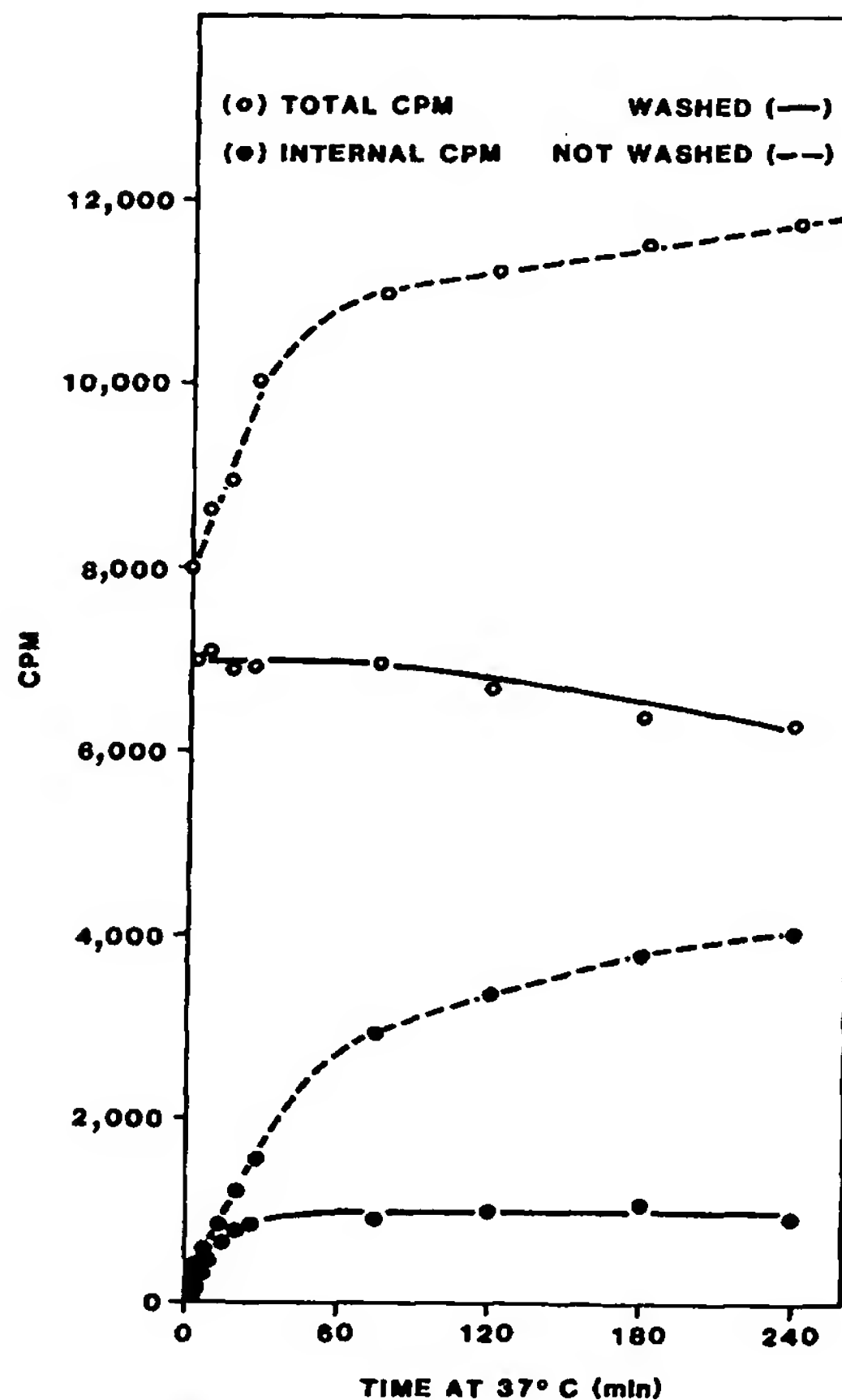


FIGURE 8. Evidence that internalization of  $^{125}\text{I}$ -IL-1 occurs via high-affinity receptors. EL4-6.1 cells were incubated at  $4^\circ\text{C}$  in the presence of  $100\text{ pM}$   $^{125}\text{I}$ -IL-1. After 4 h, half of the cells were washed three times to remove unbound  $^{125}\text{I}$ -IL-1 (solid lines), whereas the remainder were not washed (dashed lines). The cells were then incubated at  $37^\circ\text{C}$ . Total cell-associated (○) and internalized  $^{125}\text{I}$ -IL-1 (●) was measured at different times using duplicate aliquots of  $10^6$  cells, as described in Materials and Methods.

surface, the data would be consistent with a receptor recycling time of 10–20 min.

Finally, Fig. 9 shows that the capacity of cells to internalize IL-1 correlates with the number of high affinity IL-1-R expressed per cell. IL-1-R $^-$  cells such as EL4-3, CTLL, and BW5147 could not rapidly internalize  $^{125}\text{I}$ -IL-1. LBRM cells, which express  $\sim 50$  high-affinity IL-1-R per cell internalized fewer molecules than did EL4-6.1 cells, which express  $\sim 200$ – $400$  such receptors per cell.

#### Discussion

We have previously described the EL4-6.1 thymoma line as a high IL-1-responder and have shown that these cells are induced in the presence of rIL-1



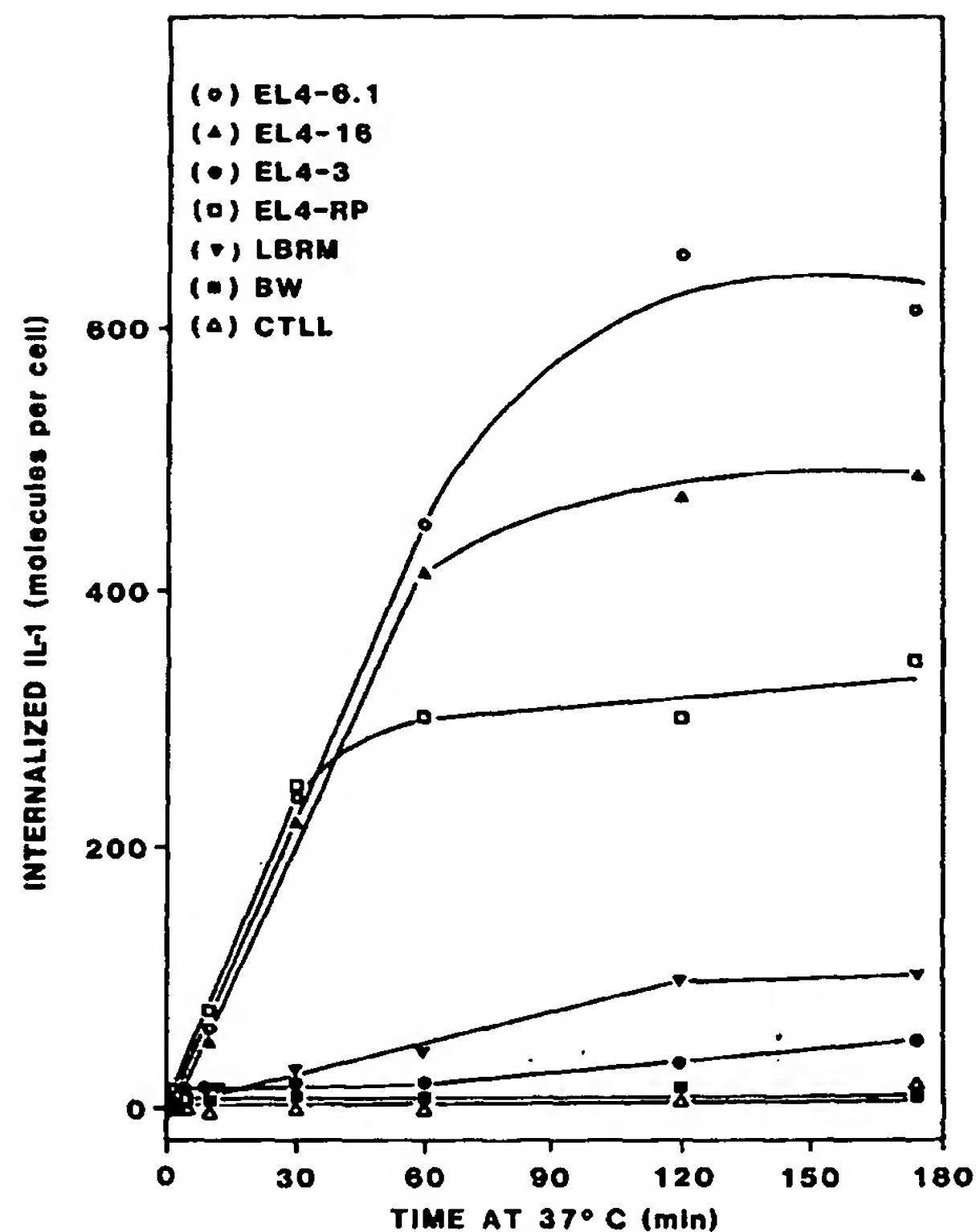


FIGURE 9. Internalization of IL-1 by various T cell lines. Cells were incubated at 37°C in the presence of 20 pM  $^{125}\text{I}$ -IL-1 and internalized  $^{125}\text{I}$ -IL-1 was measured at various times. See Table I for the number of high-affinity IL-1-R per cell.

( $\alpha$  or  $\beta$  form) to both secrete IL-2 and express IL-2-R, provided that a second signal is given in the form of Ionomycin or suboptimal concentration of PMA (13). The recent availability of human rIL-1 (23) and a procedure for its radiolabeling to high specific activity has allowed us to directly quantitate the number and affinity of IL-1-R expressed by various T cells. We show here that the IL-1-responsiveness of EL4-6.1 cells correlates with the expression of an unusually high number of IL-1 binding sites ( $\sim 20,000$  per cell), which could be resolved into high- and low-affinity classes. Despite the large variation in the overall level of IL-1-R expression between different T cell types, the ratio of the number of high- to low-affinity IL-1-R was consistent. The results presented here further suggest that the biological activity of IL-1 is mediated exclusively via interaction with the minor class of high-affinity IL-1-R.

Radioiodination of rIL-1- $\alpha$  had no detectable effect on biological activity, as measured by its ability to induce IL-2 secretion and IL-2-R expression by EL4-6.1 cells. Dower et al. (12) radiolabeled biochemically purified human IL-1- $\beta$  using the Bolton-Hunter reagent. The IL-1- $\beta$ , incubated for 1 h with the reagent, was labeled to a high specific activity ( $2-5 \times 10^{15}$  cpm/mmol) but retained only

5% of its biological activity. We radioiodinated human rIL-1- $\alpha$  by incubation in the presence of Na<sup>125</sup>I and chloramine T for 10 min, thereby achieving a high level of specific activity ( $1-2 \times 10^{16}$  cpm/mmol) while retaining full biological activity. <sup>125</sup>I-IL-1- $\alpha$  competed in an equimolar fashion with unlabeled IL-1- $\alpha$  for binding to EL4-6.1 cells. Taken together with the functional studies, these data argue that <sup>125</sup>I-IL-1 can be used as a valid estimate of the behavior of the natural form of the molecule. Furthermore, in agreement with the recent findings of Dower et al. (24), unlabeled rIL-1- $\alpha$  and rIL-1- $\beta$  both competed (albeit with slightly different affinity) for the binding of labeled rIL-1- $\alpha$  suggesting that the two forms of rIL-1 bind to the same receptor.

Scatchard plot analysis of <sup>125</sup>I-IL-1 equilibrium binding revealed the presence of two classes of IL-1-R on a number of different T cell populations. The major class, making up 95–99% of the total IL-1-R, has a  $K_d$  of 300–500 pM. The minor class of receptors has an ~100-fold higher affinity for IL-1 ( $K_d$  of 2–8 pM). The level of expression of IL-1-R among different T cell lines was variable. Certain cell lines such as ST4, BW1547, CTLL, and EL4-3 were negative, whereas a variety of independently derived EL4 lines were highly positive. Other thymoma cell lines such as Yac and LBRM expressed low, but detectable numbers of IL-1-R, which is consistent with an earlier report (12) in which the number of IL-1-R expressed by a variety of T cell types ranged from 27 to 550 per cell. In the latter study, the binding affinity and receptor number they reported for the LBRM cell line were similar to what we find here for the low-affinity class of IL-1-R. The reason why these authors failed to detect the high-affinity class of IL-1-R on LBRM cells is probably technical, i.e., they may not have been able to detect binding at sufficiently low IL-1 concentration (<10 pM) because of the lower specific activity and/or biological activity of their <sup>125</sup>I-IL-1- $\beta$  preparations as compared with the <sup>125</sup>I-IL-1- $\alpha$  preparations used in the present study.

All IL-1-R<sup>+</sup> T cell populations tested expressed both high- and low-affinity classes of IL-1-R, and the affinity of IL-1 binding to either class was comparable. The existence of high- and low-affinity receptors has been reported in a number of peptide hormone systems, including nerve growth factor (29), epidermal growth factor (30), insulin (31), platelet-derived growth factor (32), and IL-2 (27, 33). In all cases, the proportion of high-affinity receptors is low (5–15% of the total) and the difference in affinity between the two classes for ligand binding is ~50–100-fold. The biological significance of this receptor dichotomy is not known.

Interestingly, normal resting T cells expressed very low numbers of high-affinity IL-1-R (~10 per cell). This value is only an estimate of the average number of receptors per cell based on the population as a whole, and it is conceivable that only a subpopulation of peripheral T cells expresses IL-1-R. In this context, recent results from experiments using highly purified populations of lymph node T cells indicate that the L3T4<sup>+</sup> subset is positive for IL-1-R expression, whereas the Lyt-2<sup>+</sup> subset is not.<sup>2</sup> Expression of IL-1-R by resting T cells is probably constitutive, since activation by PMA and Ionomycin (which induces growth of both L3T4<sup>+</sup> and Lyt-2<sup>+</sup> subsets [34]) did not significantly

<sup>2</sup> Lowenthal, J. W., and H. R. MacDonald. Expression of interleukin-1 receptors is restricted to the L3T4<sup>+</sup> subset of mature T lymphocytes. Manuscript submitted for publication.

enhance IL-1-R expression after 2 d of culture, suggesting that there was no positive or negative selection for IL-1-R-expressing cells.

In a number of other hormone systems, the ligand-receptor complex has been shown to be internalized upon ligand binding (35, 36), and this internalization is thought to be a necessary signal for the generation of biological responsiveness (37, reviewed in 38). For IL-2, internalization has been shown to occur via the high-affinity class of IL-2-R, whereas low-affinity IL-2-R are nonfunctional in this respect (39; 40; M. Nabholz, personal communication). We now provide evidence that a similar functional dissociation exists between high- and low-affinity IL-1-R. Internalization of IL-1 is a specific process, requiring the presence of cell surface IL-1-R. The proportion of total cell-bound IL-1 that can be endocytosed by EL4-6.1 cells corresponds to the proportion of IL-1 that is bound to the high-affinity class of IL-1-R. Furthermore, when EL4-6.1 cells were incubated at 4°C in the presence of either 10 pM  $^{125}\text{I}$ -IL-1, a concentration that permits occupancy of high- but not low-affinity IL-1-R, or 1 nM, a concentration that allows occupancy of all IL-1-R, and then washed and transferred to 37°C, the rate and extent of IL-1 internalization was the same (data not shown). Finally, the capacity of different cell lines to internalize IL-1 correlated closely with the level of high-affinity IL-1-R expression. As in other hormone systems, the functional significance of ligand internalization and degradation is not known. One possibility is that this process allows dissociation of the ligand from its receptor and permits the receptor to recycle back to the cell surface and to bind more ligand. So far, no function has been attributed to low-affinity hormone receptors, but it is conceivable that they represent an external pool of nonfunctional receptors that can be rapidly recruited into the functional high-affinity pool. The transition between the two classes of receptors may involve a conformational change, for example by association with a second accessory molecule. Irrespective of its functional significance, it would appear that the ability to internalize IL-1 can be used as a sensitive, independent measure for the expression of high-affinity IL-1-R.

In contrast to EL4-6.1 cells, other T cell populations (including normal activated T cells) express low levels of IL-1-R. Even though they express only 10–60 high-affinity IL-1-R per cell they can still internalize and respond to IL-1. Similar findings of low numbers (15–300 per cell) of high affinity hormone receptors have been reported for granulocyte and macrophage growth factors (41 and F. Walker and A. W. Burgess, personal communication). It is significant that the concentration of IL-1 corresponding to half-maximal occupancy of the high-affinity class of IL-1-R (i.e., the  $K_d$  value) is identical to the concentration of IL-1 that gives half-maximal biological activity (2–8 pM). This is similar to the situation found for IL-2, where the  $K_d$  value of the high-affinity class of IL-2-R corresponds to the IL-2 concentration giving half-maximal biological activity (10–30 pM) (26, 27, 33, 37). Taken together, the data presented here suggest that the biological activity of IL-1 can be attributed exclusively to its interaction with high-affinity IL-1-R.

Because of the extremely high number of IL-1-R expressed by EL4-6.1, they constitute a valuable model system with which the structure, function, and regulation of IL-1-R can be studied. Furthermore, this cell line should facilitate

the isolation of the IL-1-R in purified form and cloning of the gene(s) that encode it.

### Summary

In this report we describe, on the basis of direct IL-1 binding assays and IL-1 internalization studies, the existence of two classes of IL-1-R on a variety of T cell types. Cells of the EL4-6.1 thymoma express large numbers (~20,000 per cell) of IL-1-R that have a  $K_d$  of ~300 pM for IL-1. Even though these receptors make up 98–99% of the total IL-1-R per cell, they appear to be nonfunctional, based on their inability to endocytose IL-1. A minor class of IL-1-R (200–400/cell) has an ~100-fold higher affinity for IL-1 ( $K_d$ , ~5 pM) and can rapidly internalize the ligand upon binding. All of the biological activity of IL-1 can be shown to occur via binding to high-affinity IL-1-R since the IL-1 concentration giving half-maximum biological activity in EL4-6.1 cells corresponds precisely to the  $K_d$  of this class of receptor. Other cell types, including normal T cells, also express both high- and low-affinity IL-1-R, but the absolute number of receptors per cell is considerably less.

We thank Drs. C. S. Henney, S. Habu, and O. Kanagawa for providing cell lines. We also thank Drs. C. S. Henney and S. K. Dower for the generous gifts of rIL-1- $\alpha$  and rIL-1- $\beta$  and for helpful advice concerning the radioiodination protocol.

*Received for publication 20 March 1986 and in revised form 16 June 1986.*

### References

1. Mizel, S. B. 1982. Interleukin 1 and T cell activation. *Immunol. Rev.* 63:51.
2. Durum, S. K., J. A. Schmidt, and J. J. Oppenheim. 1985. Interleukin-1: an immunological perspective. *Annu. Rev. Immunol.* 3:263.
3. Farrar, J. J., S. B. Mizel, J. Fuller-Farrar, W. L. Farrar, and M. L. Hilfiker. 1980. Macrophage-independent activation of helper T cells. I. Production of interleukin 2. *J. Immunol.* 125:793.
4. Larsson, E.-L., N. N. Iscove, and A. Coutinho. 1980. Two distinct factors are required for induction of T-cell growth. *Nature (Lond.)* 283:664.
5. Smith, K. A., L. B. Lachman, J. J. Oppenheim, and M. F. Favata. 1980. The functional relationships of the interleukins. *J. Exp. Med.* 151:1551.
6. Smith, K. A., K. J. Gilbride, and M. F. Favata. 1980. Lymphocyte activating factor promotes T cell growth factor production by cloned murine lymphoma cells. *Nature (Lond.)* 287:853.
7. Gillis, S., and S. B. Mizel. 1981. T-cell lymphoma model for the analysis of interleukin 1-mediated T-cell activation. *Proc. Natl. Acad. Sci. USA.* 78:1133.
8. Maizel, A. L., S. R. Mehta, R. J. Ford, and L. B. Lachman. 1981. Effect of interleukin 1 on human thymocytes and purified human T cells. *J. Exp. Med.* 153:470.
9. Kasahara, T., N. Mukaida, K. Hatake, K. Motoyoshi, T. Kawai, and K. Shiori-Nadano. 1985. Interleukin 1 (IL-1)-dependent lymphokine production by human leukemic T cell line HSB.2 subclones. *J. Immunol.* 134:1682.
10. Kaye, J., S. Gillis, S. B. Mizel, E. M. Shevach, T. R. Malek, C. A. Dinarello, L. B. Lachmann, and C. A. Janeway. 1984. Growth of a cloned helper T cell line induced by a monoclonal antibody specific for the antigen receptor: interleukin 1 is required for the expression of receptors for interleukin 2. *J. Immunol.* 133:1339.



11. Williams, J. M., D. Deloria, J. A. Hansen, C. A. Dinarello, R. Loertscher, H. M. Shapiro, and T. B. Strom. 1985. The events of primary T cell activation can be staged by use of Sepharose-bound anti-T3 (64.1) monoclonal antibody and purified interleukin 1. *J. Immunol.* 135:2249.
12. Dower, S. K., S. R. Kronheim, C. J. March, P. J. Conlon, T. P. Hopp, S. Gillis, and D. L. Urdal. 1985. Detection and characterization of high affinity plasma membrane receptors for human interleukin 1. *J. Exp. Med.* 162:501.
13. Lowenthal, J. W., J.-C. Cerottini, and H. R. MacDonald. Interleukin 1-dependent induction of both interleukin 2 secretion and interleukin 2 receptor expression by thymoma cells. *J. Immunol.* In press.
14. Zubler, R. H., F. Erard, R. K. Lees, M. Van Laer, C. Mingari, L. Moretta, and H. R. MacDonald. 1985. Mutant EL-4 thymoma cells polyclonally activate murine and human B cells via direct cell interaction. *J. Immunol.* 134:3662.
15. Conlon, P. J. 1983. A rapid biologic assay for the detection of interleukin-1. *J. Immunol.* 131:1280.
16. Gillis, S., M. Fern, W. Ou, and K. A. Smith. 1978. T cell growth factor: parameters of production and a quantitative microassay for activity. *J. Immunol.* 120:2027.
17. Culvenor, J. G., A. W. Harris, T. E. Mandel, A. Whitelaw, and E. Ferber. 1981. Alkaline phosphatase in hemopoietic tumor cell lines of the mouse: high activity in cells of the B lymphoid lineage. *J. Immunol.* 126:1974.
18. Hyman, B., and V. Stallings. 1974. Complementation patterns of Thy-1 variants and evidence that antigen loss variants "pre-exist" in the parental population. *J. Natl. Cancer Inst.* 52:429.
19. Habu, S., and K. Okumura. 1984. Cell surface antigen marking the stages of murine T cell ontogeny and its functional subsets. *Immunol. Rev.* 82:117.
20. Julius, M. H., E. Simpson, and L. A. Herzenberg. 1973. A rapid method for the isolation of functional thymus-derived murine lymphocytes. *Eur. J. Immunol.* 3:645.
21. Cerottini, J.-C., H. D. Engers, H. R. MacDonald, and K. T. Brunner. 1974. Generation of cytotoxic T lymphocytes in vitro. I. Response of normal and immune spleen cells in mixed leukocyte culture. *J. Exp. Med.* 140:703.
22. Robb, J. R., P. C. Mayer, and R. Garlick. 1985. Retention of biological activity following radioiodination of human interleukin 2: comparison with biosynthetically labelled growth factor in receptor binding assays. *J. Immunol. Methods.* 81:15.
23. March, C. J., B. Mosley, A. Larsen, D. P. Cerretti, G. Braedt, V. Price, S. Gillis, C. S. Henney, S. R. Kronheim, K. Grobstein, P. J. Conlon, T. P. Hopp, and D. Cosman. 1985. Cloning, sequence and expression of two distinct human interleukin-1 complementary DNAs. *Nature (Lond.)* 315:641.
24. Dower, S. K., S. M. Call, S. Gillis, and D. L. Urdal. 1986. Similarity between the interleukin 1 receptors on a murine T-lymphoma cell line and on a murine fibroblast cell line. *Proc. Natl. Acad. Sci. USA.* 83:1060.
25. Landegren, U. 1984. Measurement of cell number using an endogenous enzyme, hexosaminidase. Applications for the detection of lymphokines and cell surface antigens. *J. Immunol. Methods.* 67:319.
26. Robb, R. J., A. Munck, and K. A. Smith. 1981. T cell growth factor receptors. Quantitation, specificity, and biological relevance. *J. Exp. Med.* 154:1455.
27. Lowenthal, J. W., R. H. Zubler, M. Nabholz, and H. R. MacDonald. 1985. Similarities between interleukin 2 receptor number and affinity on activated B and T lymphocytes. *Nature (Lond.)* 315:669.
28. Hunter, W. M., and F. C. Greenwood. 1962. Preparation of Iodine-131 labelled human growth hormone of high specific activity. *Nature (Lond.)* 194:495.
29. Sutter, A., R. J. Riopelle, R. M. Harris-Walker, and E. M. Shooter. 1979. Nerve

- growth factor receptors. Characterization of two distinct classes of binding sites on chick embryo sensory ganglia cells. *J. Biol. Chem.* 254:5972.
30. Shoyab, M., J. E. De Larco, and G. J. Todaro. 1979. Biologically active phorbol esters specifically alter affinity of epidermal growth factor membrane receptors. *Nature (Lond.)* 219:387.
  31. Corin, R. E., and D. B. Donner. 1982. Insulin receptors convert to a higher affinity state subsequent to hormone binding. *J. Biol. Chem.* 257:104.
  32. Williams, L. T., P. M. Tremble, M. F. Lavin, and M. E. Sunday. 1984. Platelet-derived growth factor receptors from a high affinity state in membrane preparations. Kinetics and affinity cross-linking studies. *J. Biol. Chem.* 259:5287.
  33. Robb, R. J., W. C. Greene, and C. M. Rusk. 1984. Low and high affinity cellular receptors for interleukin 2. Implications for the level of Tac antigen. *J. Exp. Med.* 160:1126.
  34. Erard, F., M. Nabholz, A. Dupuy-D'Angeac, and H. R. MacDonald. 1985. Differential requirements for the induction of interleukin 2 responsiveness in L3T4<sup>+</sup> and Lyt-2<sup>+</sup> T cell subsets. *J. Exp. Med.* 162:1738.
  35. Carpentier, J.-L., P. Gordon, M. Amherdt, E. Van Obberghen, C. R. Kahn, and L. Orci. 1978. <sup>125</sup>I-Insulin binding to cultured human lymphocytes. Initial localization and fate of hormone determined by quantitative electron microscope autoradiography. *J. Clin. Invest.* 61:1057.
  36. Gordon, P., J.-L. Carpentier, S. Cohen, and L. Orci. 1978. Epidermal growth factor: morphological demonstration of binding, internalization and lysosomal association in human fibroblasts. *Proc. Natl. Acad. Sci. USA* 75:5025.
  37. Smith, K. A. 1984. Interleukin 2. *Annu. Rev. Immunol.* 2:319.
  38. Goldstein, J. L., R. G. W. Anderson, and M. S. Brown. 1979. Coated pits, coated vesicles and receptor-mediated endocytosis. *Nature (Lond.)* 279:679.
  39. Fujii, M., K. Sugamura, K. Sano, M. Nakai, K. Sugita, and Y. Hinuma. 1986. High-affinity receptor mediated internalization and degradation of interleukin 2 in human T cells. *J. Exp. Med.* 163:550.
  40. Weissman, A. M., J. B. Harford, P. B. Svetlik, W. L. Leonard, J. M. Depper, T. A. Waldmann, W. C. Greene, and R. D. Klausner. 1986. Only high affinity receptors for interleukin-2 mediate internalization of ligand. *Proc. Natl. Acad. Sci. USA* 83:1463.
  41. Nicola, N. A., and D. Metcalf. 1984. Binding of the differentiation-inducer, granulocyte-colony-stimulating factor to responsive but not unresponsive leukemic cell lines. *Proc. Natl. Acad. Sci. USA* 81:3765.

## T-Cell lymphoma model for the analysis of interleukin 1-mediated T-cell activation

(T-cell growth factor/lymphocyte-activating factor/interleukins 1 and 2/T-cell differentiation/lymphoma)

STEVEN GILLIS\* AND STEVEN B. MIZEL†‡

\*Program in Basic Immunology, Fred Hutchinson Cancer Research Center, 1124 Columbia Street, Seattle, Washington 98104; and †Laboratory of Microbiology and Immunology, National Institute of Dental Research, National Institutes of Health, Bethesda, Maryland 20014

Communicated by Lloyd J. Old, October 9, 1980

**ABSTRACT** Several laboratories have recently demonstrated that the requirement for macrophages in mitogen-induced production of murine T-cell interleukin 2 (IL-2; formerly referred to as "T-cell growth factor") could be circumvented by using the macrophage-derived peptide interleukin 1 (IL-1; formerly referred to as "lymphocyte-activating factor"). Using two cloned T-cell lymphomas, we investigated the mechanism through which IL-1 exerted its effect on IL-2 production. One of the cell lines used (LBRM-33 5A4) produces large concentrations of IL-2 upon mitogen stimulation, whereas the second (LBRM-33 1A5) is incapable of producing IL-2 in response to mitogen. It was observed that addition of purified IL-1 to nonproducer 1A5 cells converted them to a state in which subsequent mitogen stimulation triggered production of IL-2. The concentration of IL-2 produced by IL-1 treated 1A5 cells was equivalent in magnitude to that generated by mitogen-stimulated 5A4 cells (500–1000 units/ml, or approximately 1000 times the concentration of IL-2 contained in conventional preparations of murine mitogen-conditioned medium). The observations that (i) brief exposure to IL-1 was sufficient for 1A5 cell conversion to IL-2 production and (ii) IL-1 could actively be absorbed from culture medium by live or fixed 1A5 cells led us to propose the existence of IL-1 receptors on responsive 1A5 cells. On the basis of these experiments, we have postulated that IL-1 mediates its effect on immune reactivity (enhancement of thymocyte mitogenesis and induction of antibody and cytotoxic T cell responses) by maturation of a subset of immature T cells to the point where they are capable of IL-2 production. Subsequent release of IL-2 after ligand activation allows for clonal expansion of activated T cells which mediate particular effector functions.

Recent experiments conducted in several laboratories have confirmed the pivotal roles that interleukin 1 (IL-1; formerly referred to as "lymphocyte-activating factor" or LAF) and interleukin 2 (IL-2; formerly referred to as "T-cell growth factor" or TCGF) play in the generation of T and B cell immune reactivities (1–8). Both proteins have been shown (i) to enhance thymocyte mitogenesis, (ii) to support the induction of alloantigen-primed cytotoxic T-cell reactivity, and (iii) to aid in the generation of helper T cells for antibody responses after stimulation with heterologous erythrocytes (1, 2, 4–7). In marked contrast, IL-2 is the sole interleukin capable of sustaining the *in vitro* exponential proliferation of effector T-cell lines (1, 2, 8, 9). Similarly, only IL-2 has been shown to be capable of allowing for *in vitro* and *in vivo* generation of cytotoxic T cells from *nude* mouse spleens (1, 2, 10, 11).

Due to the well-documented macrophage requirement for mitogen-stimulated IL-2 production (12), and the observation that IL-1 is a macrophage product (13), it has been hypothesized that IL-1 may be an essential signal required by IL-2 producer T cells. Such a hypothesis was first suggested by the observations of Smith *et al.* (12) who found that IL-1 producer tumor cell

line supernates could restore adherent cell-depleted, mitogen-stimulated, murine spleen cell cultures to normal levels of IL-2 production. IL-1 involvement in IL-2 production was also suggested by the studies of Farrar *et al.* (14) who showed that the tumor promotor phorbol myristate acetate [PMA; previously shown to act as a replacement for IL-1 in several immune response assays (15, 16)] could substitute for the macrophage requirement for IL-2 production by murine T cells. Finally, addition of IL-1 to mitogen-activated, adherent, cell-depleted, spleen cell cultures has been shown to reconstitute IL-2 production to normal levels (17, 18). Although the involvement of IL-1 in the production of IL-2 by ligand-stimulated T cells cannot be questioned, a precise mechanism by which IL-1 functions in this capacity has not been presented. For example, it is not known if IL-1 and mitogen affect the same cell or different cells that then must interact to initiate IL-2 production.

Experiments in our laboratories have recently documented the existence of an extremely potent IL-2 producer lymphoma cell, LBRM-33 (19). Mitogen-stimulated LBRM-33 cells produce 1000–10,000 times the amount of IL-2 generated by conventional cultures of mitogen-stimulated mouse spleen cells. Limiting-dilution cloning of LBRM-33 resulted in the isolation of both extremely high titer IL-2 producer clones and one nonproducer lymphoma cell line variant (19). On the basis of these observations, we initiated studies to evaluate the utility of IL-2 producer and nonproducer LBRM-33 clones in the dissection of the IL-1 requirement for IL-2 production.

In this communication we report that IL-1 has the capacity to convert IL-2 nonproducer tumor clones to high-titer IL-2 production. Conversion to IL-2 production occurs after brief exposure to IL-1 at concentrations that do not produce detectable thymocyte mitogenesis (13), thereby leading to the development of a more sensitive bioassay for IL-1 activity. Furthermore, the observation that LBRM-33 cells could absorb IL-1 from cultures suggests that these cells possess surface receptors for IL-1 and identifies a valuable cell population for further study of the molecular aspects of IL-1-induced immune reactivity.

### MATERIALS AND METHODS

**Cell Lines.** LBRM-33, a radiation-induced splenic lymphoma from the B10.BR mouse (originally isolated by G. Cudkowicz, Roswell Park Memorial Institute, Buffalo, NY), has been shown (19) to produce high-titer murine IL-2 upon 24-hr stimulation with T-cell mitogens. Limiting-dilution cloning of

Abbreviations: IL-1, interleukin 1; IL-2, interleukin 2; PHA, phytohemagglutinin; PMA, phorbol myristate acetate.

‡ Present address: Department of Microbiology, Cell Biology, Biochemistry and Biophysics, The Pennsylvania State University, University Park, PA 16802.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.



LBRM-33 resulted in identification of both extremely high titer IL-2 producer (clone 5A4, 1000 units of IL-2 per ml) and non-producer cell line (LBRM-33 1A5) clones. LBRM cells were further found to be positive for cytosol terminal deoxynucleotidyl transferase and expressed TL, Thy 1, Lyt 1, Lyt 2, and Lyt 3 markers at the cell surface. 5A4 and 1A5 cells used in these studies were maintained *in vitro* in RPMI-1640 medium supplemented with 5% heat-inactivated (56°C, 30 min) fetal calf serum, 50  $\mu$ M 2-mercaptoethanol, penicillin (50 units/ml), streptomycin (50  $\mu$ g/ml), and fresh L-glutamine (300  $\mu$ g/ml). IL-2 production cultures were conducted in this medium in either 200  $\mu$ l (no. 3596, flat-bottom microplate, Costar, Cambridge, MA) or 5 ml (no. 3013 tissue culture flasks, Falcon Plastics, Oxnard, CA). IL-2 production was initiated by addition of phytohemagglutinin M (PHA; 0.1–1% by volume, GIBCO) to LBRM cell line cultures ( $10^6$  cells per ml). In several experiments, 1A5 and 5A4 cells were also stimulated with either PMA (10 ng/ml, Sigma) or IL-1 (for concentrations see *Results*) prepared as detailed below. Supernatants harvested from 24-hr cultures were tested for IL-2 activity as detailed below. Maintenance of LBRM-33 cell lines and IL-2 production experiments were at 37°C in a humidified atmosphere of 5% CO<sub>2</sub> in air.

**IL-2 Assay.** The 24-hr culture supernatants from stimulated LBRM-33 5A4 and 1A5 cells (either in the presence or in the absence of IL-1 at various concentrations) were tested for IL-2 activity by using a standard microassay (20) based on the IL-2-dependent exponential proliferation of a murine cytotoxic T-cell line (CTLL) (8). Briefly, 3000 CTLL cells were cultured in replicate 200- $\mu$ l volumes in flat-bottomed microplate wells in the presence of a log<sub>2</sub> dilution series of putative IL-2-containing samples. After 24 hr, the cells were exposed to 0.5  $\mu$ Ci (1 Ci =  $3.7 \times 10^{10}$  becquerels) [<sup>3</sup>H]dThd (20 mCi/mmol; New England Nuclear) for an additional 4 hr after which the cultures were harvested onto glass fiber filter strips with the aid of a multiple automated sample harvester (MASH 11, Microbiological Associates, Bethesda, MD). [<sup>3</sup>H]dThd incorporation was then determined by liquid scintillation counting. Only CTLL cells cultured in the presence of IL-2 incorporated [<sup>3</sup>H]dThd in a dose-dependent manner. Consistent with the observation that CTLL cells cultured in the absence of IL-2 are >95% trypan blue-positive, cultures lacking IL-2 incorporated <100 cpm of [<sup>3</sup>H]dThd. IL-2 activity was quantified by probit analysis of [<sup>3</sup>H]dThd incorporation data as described (20). A standard preparation of IL-2-conditioned medium [1 unit/ml, 48-hr supernate of concanavalin A-stimulated (5  $\mu$ g/ml) rat spleen cells ( $10^6$  cells per ml)] routinely generated 10,000–15,000 cpm of [<sup>3</sup>H]dThd incorporation at a dilution of 1:2. Similarly, a supernatant containing 1000 units of IL-2 activity per ml stimulated identical levels of [<sup>3</sup>H]dThd incorporation at a dilution of 1:2000. CTLL cells were routinely maintained *in vitro* in exponential proliferation in RPMI-1640 medium supplemented with 50% rat spleen cell-conditioned medium (produced as detailed above) and 2% fetal calf serum (37°C, 5% CO<sub>2</sub> in air).

**Preparation of IL-1 and Assay for Activity.** The IL-1 used for induction of LBRM-33 IL-2 production was prepared from the culture supernatant of PMA-stimulated P388D<sub>1</sub> macrophage tumor cells as detailed (13). IL-1 was partially purified by a sequence of differential ammonium sulfate precipitations, DEAE-cellulose ion exchange chromatography, and Sephacryl S200 gel exclusion chromatography (21). After this, all of the biologically active IL-1 was associated with a protein of  $M_r \approx 15,000$  (21). The IL-1 used in absorption experiments was further purified by phenyl-Sepharose hydrophobic affinity chromatography (22). After Sephacryl S200 gel filtration, the specific activity of the partially purified IL-1 was approximately

10,000 units/mg of protein. Phenyl-Sepharose chromatography increased the specific activity  $\approx 30$  fold.

In some experiments, IL-1 was inactivated by the arginine-modifying agent phenylglyoxal (22). IL-1 was dialyzed into 200 mM imidazole buffer (pH 8) and incubated in 1% phenylglyoxal for 6 hr at 25°C. After this treatment, IL-1 was passed over a column of Sephadex G-25 (equilibrated in phosphate-buffered saline, pH 7.2) to remove the phenylglyoxal. Control IL-1 was mixed with 1% phenylglyoxal immediately prior to gel filtration chromatography. Phenylglyoxal treatment had no effect on mobility of IL-1 on Sephadex G-25 chromatography columns.

IL-1 activity was determined by its capacity to enhance thymocyte proliferation in response to *in vitro* stimulation with PHA at 1  $\mu$ g/ml as detailed (13). Briefly, thymocytes from C3H/HeJ mice (4–6 weeks old, Jackson Laboratory, Bar Harbor, ME) were cultured in 200  $\mu$ l in flat-bottom microplate wells ( $1.5 \times 10^6$  cells per well) in the presence of a log<sub>2</sub> dilution series of putative IL-1-containing samples. Under these conditions, a standard preparation of IL-1 (100 units/ml) induced approximately 10,000–15,000 cpm of thymocyte [<sup>3</sup>H]dThd incorporation at a dilution of 1:4. Half-maximal thymocyte proliferation was routinely observed after culture stimulation with 5–10 units of IL-1 per ml. Units of IL-1 activity were quantified by probit analysis as detailed elsewhere (20).

**IL-1 Absorption.** To test the capacity of LBRM-33 cell line derivatives to absorb IL-1, both 5A4 and 1A5 cells ( $10^6$  cells) were washed and resuspended in RPMI-1640/2% fetal calf serum containing either 100 or 0.5 unit of IL-1 per ml. After a 4-hr incubation at 47 or 37°C, the cells were pelleted by a 10-min centrifugation at  $300 \times g$  and the supernatant was tested for residual IL-1 activity by its capacity to enhance thymocyte mitogenesis or to induce LBRM-33 1A5 IL-2 production. In some experiments, LBRM-33 5A4 and 1A5 cells were fixed with 2% glutaraldehyde (15 min, 4°C) and washed five times with 50 ml of culture medium prior to use in absorption tests.

## RESULTS

**Tumor Cell Line IL-2 Production.** The results presented in Table 1 review the relative capacities of LBRM-33 5A4 and 1A5 clones to produce IL-2 after incubation with 1% PHA. As previously observed (19), 5A4 cells cultured for 24 hr ( $10^6$  cells per ml) in the presence of 1% PHA produced approximately 1000 units of IL-2 activity per ml whereas LBRM-33 1A5 cells produced no detectable IL-2. It should be stressed that the amount of IL-2 generated by mitogen-stimulated 5A4 cultures was be-

Table 1. IL-2 production by LBRM-33 5A4 and 1A5 cell lines

Stimulated with*			IL-2 present in 24-hr supernate, units/ml
1% PHA	0.1% PHA	PMA (10 ng/ml)	
With LBRM-33 5A4†			
+	—	—	813
—	+	—	26
—	—	+	0
+	—	+	869
—	+	+	762
With LBRM-33 1A5†			
+	—	—	0
—	+	—	0
—	—	+	0
+	—	+	32
—	+	+	12

\* +, Present in culture; —, absent from culture.

†  $10^6$  cells per ml in RPMI-1640/2% fetal calf serum.



tween 1000 and 10,000 times the amount of IL-2 produced by identical concentrations of normal murine spleen cells. We routinely found 1% PHA to be an optimal mitogen dose for eliciting production of IL-2 by the LBRM-33 5A4 cell line. Reduction of this mitogen concentration by a factor of 10 (0.1% PHA) significantly diminished the amount of IL-2 produced by clone 5A4; stimulation of nonproducer 1A5 cells with 0.1% PHA did not lead to IL-2 production.

It was interesting that incubation of 5A4 cells with 0.1% PHA and PMA at 10 ng/ml resulted in maximal IL-2 production. Of potential importance was the observation that PMA in the presence of PHA stimulated comparatively weak but significant, IL-2 production by 1A5 cells. Therefore, it was conceivable that 1A5 cells possessed the capacity to secrete IL-2 but apparently required additional signals to activate the IL-2 production process. Several laboratories have demonstrated an adherent cell requirement for murine IL-2 production (12, 14, 17, 18). In the absence of macrophages, concanavalin A-stimulated, purified T-cell populations could be restored to normal levels of IL-2 production by addition of either macrophage tumor cell line supernatants (12) or the monokine IL-1 (17, 18). Farrar *et al.* (14) found that addition of PMA to mitogen-stimulated, macrophage-depleted, spleen cell cultures also reconstituted normal IL-2 production. These results coupled with those detailed in Table 1 led us to examine the effect of IL-1 on both the PHA-responsive LBRM-33 5A4 (IL-2 producer) and the PHA-unresponsive LBRM-33 1A5 (IL-2 nonproducer) cloned lymphoma cell lines.

**Effect of IL-1 on Production of IL-2 by Tumor Cell Line.** Addition of P388D<sub>1</sub> macrophage-derived IL-1 to cultures of PHA-stimulated LBRM-33 5A4 and 1A5 cells had profound effects on IL-2 production. As with PMA, addition of IL-1 at 10 units/ml to suboptimally mitogen-stimulated 5A4 cells resulted in restoration of peak levels of IL-2 production (Table 2). Furthermore, IL-1 at 10 units/ml also induced optimal IL-2 production by PHA-stimulated 1A5 cultures. IL-1 induced high-titer IL-2 production by 1A5 cells when the cells were stimulated with optimal (1%) or suboptimal (0.1%) concentrations of PHA. Addition of IL-1 to cultures of non-mitogen-stimulated 1A5 cells was not a sufficient stimulus for conversion of the 1A5 cell line to IL-2 production.

Perhaps the most striking aspect of the capacity of IL-1 to foster 1A5 cell IL-2 production was the extremely low concentrations of IL-1 at which such a conversion was observed. Dose-response curves for both IL-1-induced murine thymocyte proliferation and IL-1-dependent conversion of 1% PHA-stimulated 1A5 cells to IL-2 production are shown in Fig. 1.

Table 2. Effect of IL-1 on IL-2 production by LBRM-33 5A4 and 1A5 cell lines

Culture*	IL-2 in 24-hr supernate, units/ml
5A4 + 1% PHA	565
5A4 + 0.1% PHA	35
5A4 + IL-1	0
5A4 + IL-1 + 1.0% PHA	604
5A4 + IL-1 + 0.1% PHA	525
1A5 + 1% PHA	0
1A5 + 0.1% PHA	0
1A5 + IL-1	0
1A5 + IL-1 + 1% PHA	476
1A5 + IL-1 + 0.1% PHA	513

\* Cultures contained  $10^6$  cells per ml in RPMI-1640/2% fetal calf serum. When present, IL-1 was at 10 units/ml.

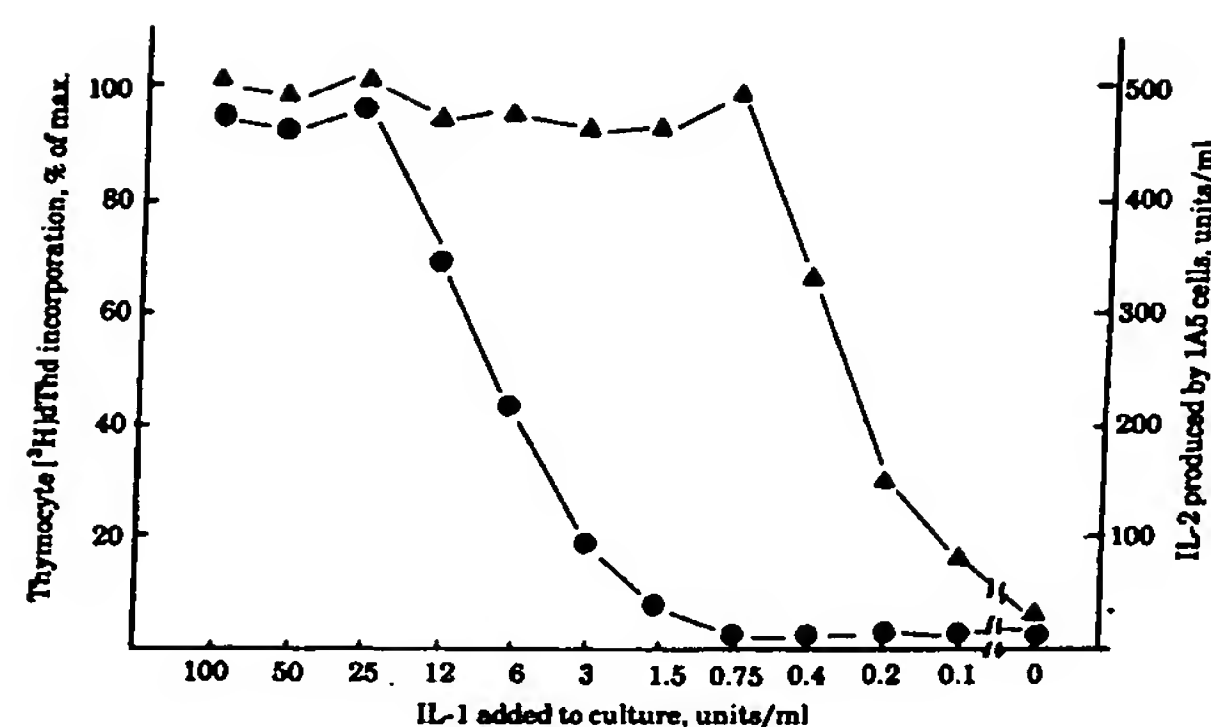


FIG. 1. Dose-response of IL-1 activity as measured by induction of thymocyte mitogenesis (●) and IL-2 production by LBRM-33 1A5 cells (▲).

Fifty percent of maximal IL-1-induced thymocyte proliferation was observed at an IL-1 concentration of approximately 5 units/ml. However, IL-1 at 5 units/ml stimulated 1A5 cells to peak levels of IL-2 production. The capacity of IL-1 to convert 1A5 cells to maximum production of IL-2 did not begin to decline until IL-1 concentrations as low as 0.5 units/ml were tested. In fact, PHA-stimulated 1A5 cells cultured in the presence of IL-1 at 0.2 unit/ml produced a significant amount of IL-2 (165 units/ml). The observation that normal T-cell mitogen-stimulated murine spleen cells produce, under optimal conditions ( $10^7$  cells per ml; concanavalin A at  $2.5 \mu\text{g/ml}$ ) only 0.5–1.5 units of IL-2 per ml further substantiates the capacity of low concentrations of IL-1 to convert nonproducer 1A5 cells to high-titer IL-2 production.

We are confident that the effects of IL-1 shown in Fig. 1 were due to the monokine itself as opposed to some unidentified contaminant molecule present in the IL-1 preparation. The IL-1 used in these studies was generated in a four-step purification procedure and was of relatively high specific activity. Additionally, treatment of purified IL-1 with 1% phenylglyoxal totally eliminated not only its capacity to augment thymocyte mitogenesis but also its ability to convert 1A5 cells to IL-2 production.

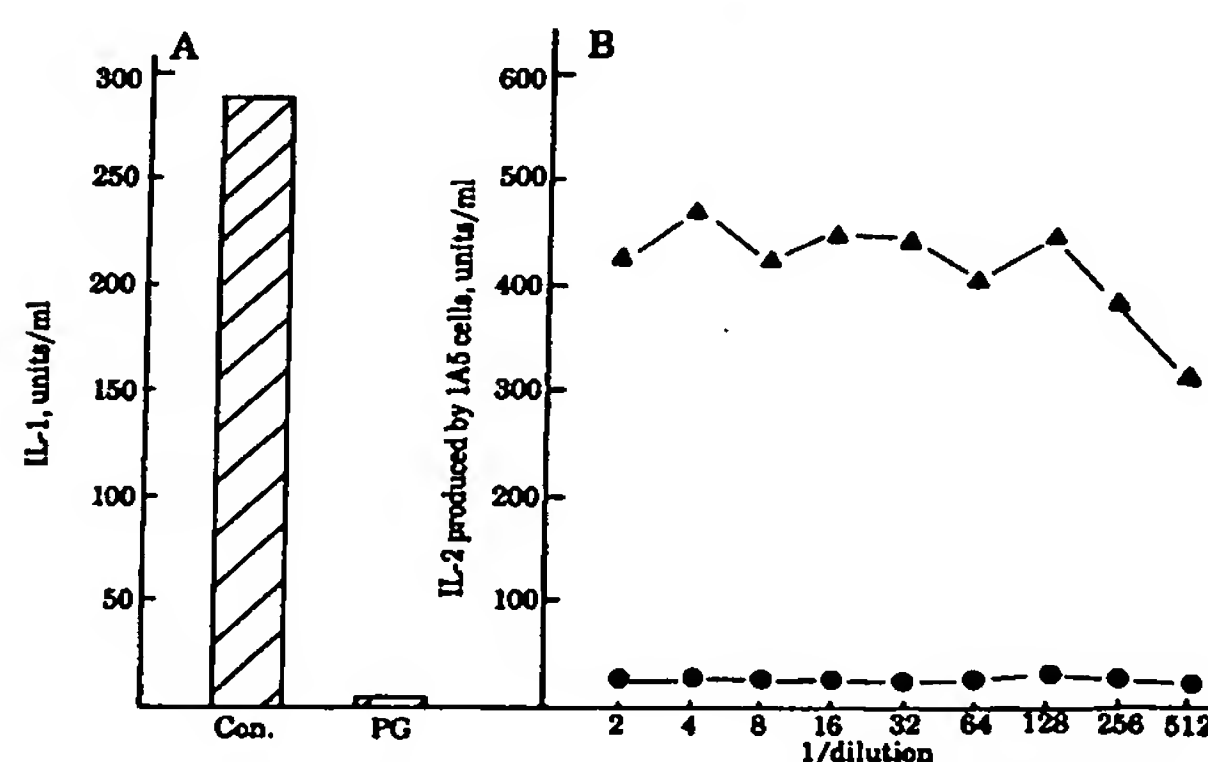


FIG. 2. Phenylglyoxal-mediated inhibition of IL-1 activity. IL-1 (300 units/ml) was exposed to 1% phenylglyoxal for 6 hr ( $25^\circ\text{C}$ ) and then was separated from the phenylglyoxal by Sephadex G-25 gel exclusion chromatography. (A) Phenylglyoxal (PG) treatment abolished all IL-1 activity as measured by its capacity to augment thymocyte mitogenesis. (B) Capacity of a log<sub>2</sub> dilution series of both control (▲) and phenylglyoxal-treated (●) IL-1 to convert PHA-stimulated 1A5 cells to IL-2 production.

(Fig. 2). Previous studies (22) have shown that phenylglyoxal modification of arginine residues in IL-1 results in destruction of its biologic activity in a number of different biological systems (without altering its mobility on gel exclusion chromatography columns). As a site-specific modifier, it seems unlikely that phenylglyoxal treatment (and destruction of conventionally tested IL-1 activity) would have a similar effect on some other contaminant protein present in the IL-1 population which was responsible for converting 1A5 cells to IL-2 production.

**Absorption of IL-1 Activity by LBRM-33 5A4 and 1A5 Cell Lines.** It was clear that IL-1 itself was not a sufficient stimulus for IL-2 production in that 1A5 cells exposed to IL-1 in the absence of mitogen did not produce IL-2. In an attempt to dissociate requirements for IL-1 and mitogen sensitization, we tested whether conversion of 1A5 cells to a state in which they were capable of producing IL-2 required the continued presence of IL-1. Multiple cultures were prepared in which 1A5 cells were either treated continuously with both PHA and IL-1 (10 units/ml) or sequentially with IL-1 and then PHA. 1A5 cells were first exposed to IL-1 for 4 hr at 37°C; after exhaustive washing, the cells were cultured in either the presence or absence of additional IL-1 and PHA. Brief exposure of 1A5 cells to IL-1 did not affect their capacity to produce maximal levels of IL-2 upon subsequent exposure to PHA (Table 3).

These results suggested that the interaction of IL-1 with 1A5 cells was relatively rapid in nature. Based on previous studies (3) that demonstrated the capacity of activated T cells to absorb IL-2 from cultures, we questioned whether 1A5 or 5A4 cells possessed a similar cell surface responsiveness (presumably mediated by receptors) for IL-1. To test this hypothesis, large numbers of 1A5 and 5A4 cells were harvested from cultures, washed extensively, and resuspended in the presence of a known amount of IL-1. After a 4-hr incubation at 4°C, the cells were pelleted and the supernatants were tested for residual IL-1 activity (assayed by enhancement of thymocyte mitogenesis as well as by conversion of 1A5 cells to IL-2 production). Absorptions were conducted with two different IL-1 concentrations: 200 units/ml in experiments in which activity was to be tested by conventional IL-1 assay (thymocyte proliferation); and 0.5 unit/ml in experiments in which absorbed supernatants were assayed by their effect on 1A5 cells. Different concentrations of IL-1 were used in the two assay systems because of the significant difference in IL-1 dose-responses noted in the two bioassay systems (Fig. 1).

Both 1A5 and 5A4 cell lines were capable of absorbing IL-1 activity (Fig. 3). As measured by the ability of the residual supernatant IL-1 to enhance thymocyte mitogenesis, 1A5 cells possessed a greater capacity on a per cell basis to absorb IL-1. Perhaps this was predictable from the observation that IL-1 converted the PHA-unresponsive 1A5 cell line (an IL-2 nonproducer) to IL-2 production whereas its effect on a similar cell line already capable of IL-2 production (5A4) was simply to enhance the effect of PHA. We are confident that the absorption shown

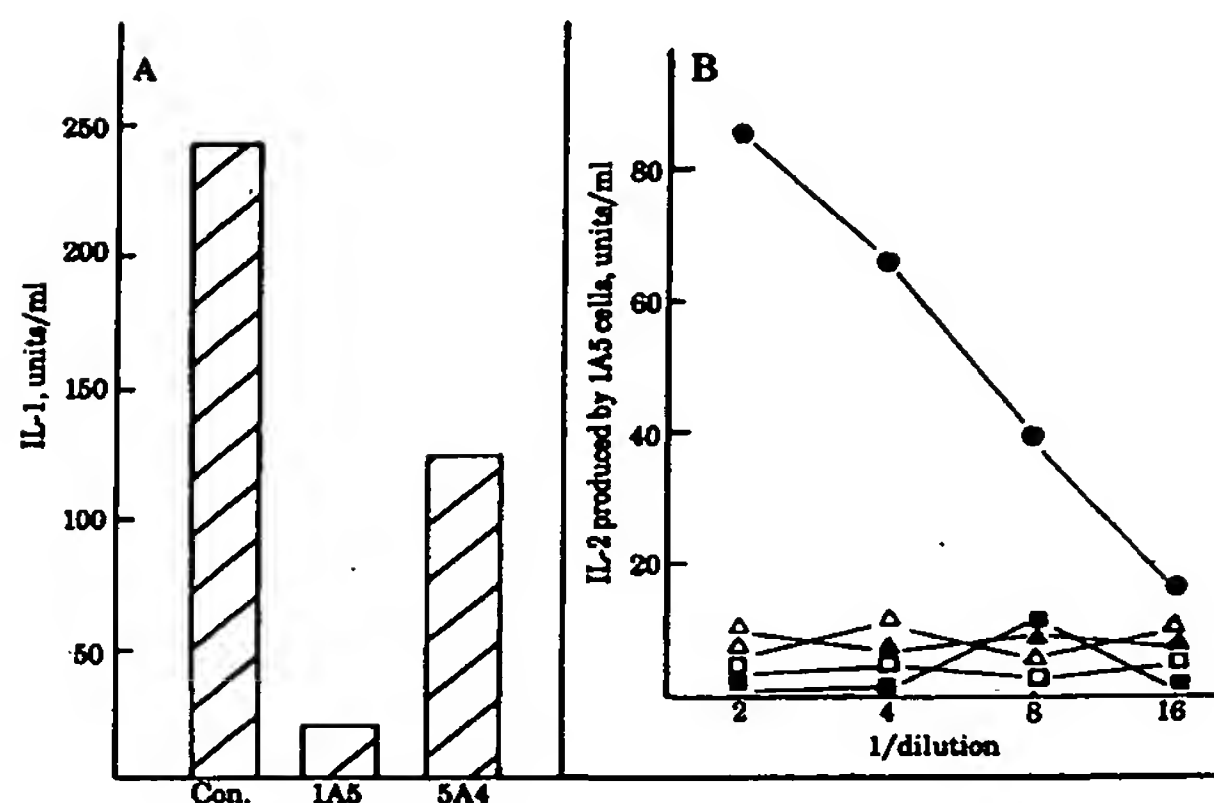


FIG. 3. Absorption of IL-1 by 1A5 and 5A4 cells. (A) IL-1 (242 units/ml) was incubated for 4 hr at 4°C either alone (control) or in the presence of  $10^6$  1A5 or 5A4 cells. Then the cells were removed and the supernatants were tested for residual IL-1 activity by their ability to enhance thymocyte mitogenesis. (B) Capacity of these samples to induce IL-2 production by 1A5. ●, Control IL-1 (0.5 unit/ml); ▲, IL-1 (0.5 unit/ml) after absorption with live or 2% glutaraldehyde-fixed 1A5 cells; ■, IL-1 (0.5 unit/ml) after absorption with live or 2% glutaraldehyde-fixed 5A4 cells.

in Fig. 3 was due to expression of a cell-surface responsiveness to IL-1 rather than to an LBRM-33 cell-mediated degradation of IL-1 activity because: (i) absorptions were conducted at 4°C, a condition not supportive of high levels of cellular metabolism, and (ii) identical absorption characteristics were found when trials were conducted with 2% glutaraldehyde-fixed 1A5 and 5A4 cells (Fig. 3).

## DISCUSSION

The experiments presented in this report took advantage of the existence of two cloned T-cell lymphomas of identical lineage (19). One cell line (LBRM-33 5A4) produces large quantities of IL-2 (the interleukin responsible for the sustained proliferation of activated T lymphocytes) whereas the second (LBRM-33 1A5) is incapable of IL-2 production after PHA stimulation. Because several previous investigations have suggested a requirement for the monokine IL-1 in the production of IL-2 by normal murine splenic T cells (12, 14, 17, 18), we used these two lymphoma cell lines in an attempt to dissect the mechanism(s) by which IL-1 participates in the initiation of IL-2 production.

We found that IL-1 has the capacity to modify 1A5 cells in some manner to render them capable of producing IL-2 upon subsequent exposure to a T-cell mitogen (PHA). Such activation occurs after brief exposure to IL-1 at a concentration approximately 1/10th that shown to be effective in the conventional IL-1 assay based on enhancement of thymocyte mitogenesis. This observation is not totally surprising if one considers (on the basis of the experiments detailed in this report) what is truly being measured in a conventional IL-1 assay. It has been well documented that IL-1 *per se* is incapable of triggering the sustained proliferation of activated T lymphocytes (6). This is a property unique to IL-2. Therefore, the enhanced mitogenesis observed when an entire thymocyte preparation is stimulated with PHA and IL-1 must be dependent upon production of the mitogenic factor IL-2. Because only the relatively small, cortisol-resistant population of the thymus (i.e., the mature T-cell compartment) has been shown to produce IL-2 (12), stimulation of an entire thymocyte preparation with mitogen (in the absence of exogenous IL-1) results in poor IL-2 production and concomitant T-

Table 3. Conversion of 1A5 cells to IL-2 production does not require continual presence of IL-1

Pretreatment (4 hr)	Subsequent 24-hr culture	IL-2 in 24-hr supernate, units/ml
None	1% PHA	0
None	1% PHA + IL-1	379
IL-1	1% PHA	410
IL-1	1% PHA + IL-1	426

When present, IL-1 was at 10 units/ml. The cells were cultured in RPMI-1640/10% fetal calf serum.



cell proliferation. However in the presence of IL-1, presumably more T cells (perhaps immature thymocytes) are converted to a state in which they can produce IL-2 and, in so doing, markedly augment thymocyte mitogenesis.

Based on such a mechanism of action for IL-1, it is not at all surprising that IL-1 converts 1A5 cells to IL-2 production at a far lower concentration than has been shown to be necessary for thymocyte proliferation. Conversion to IL-2 production as measured on 1A5 cells focuses the effects of IL-1 on a homogeneous cloned cell population whereas the percentage of IL-1-responsive T cells present in an entire thymus preparation is more than likely diminished by the presence of several if not hundreds of distinct T-cell subpopulations at varying levels of IL-1 responsiveness. It is unknown at present which cells in the thymus respond to IL-1 and if those cells (as is the case with 1A5) respond by maturing to a state of potential IL-2 production. However, by using sensitive assay systems which can now be developed due to IL-1 effects on 1A5 cells, it may be possible to isolate and to identify this naturally occurring, IL-1-responsive, thymocyte subset.

To this end, the observation that 1A5 cells and, to a lesser extent, 5A4 cells possess cell surface receptors to IL-1 is extremely important. 1A5 cells were shown to be converted to a state of lectin-inducible IL-2 production after only brief exposure to IL-1. These same 1A5 cells were also shown to be capable of absorbing IL-1. In fact, both live and fixed 1A5 cells exhibited identical absorption of IL-1 activity in 4 hr incubations at 4°C. These results argue in favor of the existence of a 1A5 cell surface responsiveness to IL-1 (presumably via the cell surface receptors). Further evidence in support of the existence of 1A5 cell surface IL-1 receptors comes from recent experiments conducted in our laboratories. Preliminary evidence indicates that biochemically identifiable IL-1 can be eluted from fixed 1A5 cells by heating the cells to 56°C or by washing the 1A5 cell pellet with 0.1% acetic acid. When the material extracted by heat or acid was lyophilized and electrophoresed on two-dimensional Tris-glycinate polyacrylamide gels, silver-stainable protein bands were detected that were identical in charge and molecular weight to purified IL-1 (unpublished observations).

It should be stressed that previous attempts to use whole thymocyte populations to absorb IL-1 activity consistently have been unsuccessful. Such results might be due to the relatively small percentage of IL-1 responsive cells in the thymus that are converted to a state of potential IL-2 production. It is hoped that, as radiolabeled, purified IL-1 becomes available, 1A5 cells will prove to be a valuable reagent in standardizing assays for IL-1 surface binding and perhaps quantifying numbers of IL-1 cell surface receptors. With such assay systems at hand it would be possible to probe various T cell populations (selected on the basis of either cell surface or functional phenotype) systematically for responsiveness to IL-1 and conversion to IL-2 production.

In any event, it is clear that 1A5 cells will be extremely useful in probing the molecular effects of IL-1 responses. The availability of a cloned cell line that responds after brief exposure to purified monokine allows study of the events that program IL-2 production at the level of gene activation, RNA transcription,

and protein synthesis. Additionally, 1A5 cells will be extremely useful in determining the consequences of IL-1 responses in terms of alteration of cell surface phenotype and function. In this regard, it is interesting to note that 1A5 cells contain high levels of cytosol terminal deoxynucleotidyltransferase (unpublished observations). In that this enzyme is a biochemical marker typical of immature T cells, such a 1A5 cell phenotype is (i) suggestive that naturally occurring IL-1-responsive cells might be found in the immature compartment of the thymus and (ii) predictive that IL-1 should have little effect on mature T cells lacking this transferase (a contention that is well supported by present studies detailing IL-1 action). Finally, by converting 1A5 cells to a state of potential IL-2 production, it appears that the monokine achieves its effect on immune responses via a differentiation/maturation-inducing pathway as opposed to triggering cellular proliferation (an activity previously shown to be associated with IL-2).

S.G. is a Special Fellow of the Leukemia Society of America and is supported in part by National Cancer Institute Grant CA28419 and Grant 1-724 from the National Foundation-March of Dimes.

1. Watson, J., Gillis, S., Marbrook, J., Mochizuki, D. & Smith, K. A. (1979) *J. Exp. Med.* 150, 849-861.
2. Gillis, S., Smith, K. A. & Watson, J. (1980) *J. Immunol.* 124, 1954-1962.
3. Smith, K. A., Gillis, S., Ruscetti, F. W., Baker, P. E. & McKenzie, D. (1979) *Proc. N.Y. Acad. Sci.* 332, 423-432.
4. Farrar, J. J., Simon, P. L., Koopman, W. J. & Fuller-Bonar, J. (1978) *J. Immunol.* 121, 1353-1360.
5. Simon, P. L., Farrar, J. J. & Kind, P. D. (1979) *J. Immunol.* 122, 127-132.
6. Aarden, L. A. *et al.* (1979) *J. Immunol.* 123, 2928-2929.
7. Koopman, W. J., Farrar, J. J. & Fuller-Bonar, J. (1978) *Cell Immunol.* 35, 92-98.
8. Gillis, S. & Smith, K. A. (1977) *Nature (London)* 268, 154-156.
9. Gillis, S., Baker, P. E., Ruscetti, F. W. & Smith, K. A. (1978) *J. Exp. Med.* 148, 1093-1098.
10. Gillis, S., Union, N. A., Baker, P. E. & Smith, K. A. (1979) *J. Exp. Med.* 149, 1460-1476.
11. Wagner, H., Hardt, C., Heeg, K., Rollinghoff, M. & Pfizenmaier, K. (1980) *Nature (London)* 284, 278-280.
12. Smith, K. A., Gillis, S. & Baker, P. E. (1979) in *The Molecular Basis of Immune Cell Function*, ed. Kaplan, J. G. (Elsevier/North Holland, Amsterdam), pp. 223-231.
13. Mizel, S. B., Oppenheim, J. J. & Rosenstreich, D. L. (1978) *J. Immunol.* 120, 1497-1503.
14. Farrar, J. J., Mizel, S. B., Fuller-Farrar, J., Farrar, W. L. & Hilfiker, M. L. (1980) *J. Immunol.* 125, 793-798.
15. Rosenstreich, D. L. & Mizel, S. B. (1979) *J. Immunol.* 123, 1749-1754.
16. Mastro, A. M. & Mueller, G. C. (1974) *Exp. Cell Res.* 88, 40-46.
17. Larsson, E. L., Iscove, N. N. & Coutinho, A. (1980) *Nature (London)* 283, 664-666.
18. Smith, K. A., Lachman, L. B., Oppenheim, J. J. & Favata, M. F. (1980) *J. Exp. Med.* 151, 1551-1556.
19. Gillis, S., Scheid, M. & Watson, J. (1980) *J. Immunol.* 125, 2570.
20. Gillis, S., Ferm, M. M., Ou, W. & Smith, K. A. (1978) *J. Immunol.* 120, 2027-2032.
21. Mizel, S. B. (1979) *Proc. N.Y. Acad. Sci.* 332, 539-549.
22. Mizel, S. B. (1980) *Mol. Immunol.* 17, 571-577.

## PURIFICATION TO HOMOGENEITY OF B CELL STIMULATING FACTOR

### A Molecule That Stimulates Proliferation of Multiple Lymphokine-dependent Cell Lines

BY KENNETH GRABSTEIN, JUNE EISENMAN, DIANE MOCHIZUKI,  
KURT SHANEBECK, PAUL CONLON, THOMAS HOPP,  
CARL MARCH, AND STEVEN GILLIS

*From the Immunex Corporation, Seattle, Washington 98101*

The differentiation of resting B lymphocytes to Ig secretion involves several sequential steps regulated by antigen, T lymphocytes, and macrophages. The requirement for T cells may in some cases be replaced by lymphokines secreted upon T cell activation (1, 2). A number of distinct T cell-derived lymphokines have been described (3-9) that directly regulate the growth and maturation of B cells. These include IL-2 (3, 4), IFN- $\gamma$  (5, 6), and at least two additional, yet less well-characterized molecules, B cell stimulating factor, (also known as B cell growth factor, BCGF or BSF-1)<sup>1</sup> (7, 8), and B cell differentiation factor (BCGF-II or BCDF) (9). BSF-1 was originally (7) described as a factor that stimulated proliferation of B cells in conjunction with a submitogenic concentration of anti-Ig. Recent studies (10, 11) have shown that semipurified BSF-1 acts on resting B cells, facilitating their entry into S phase upon subsequent interaction with anti-Ig. These studies (11) indicate that BSF-1 might actually be a differentiation factor, and they further suggest that BSF-1 may not have growth factor activity.

We have purified BSF-1 to homogeneity from culture supernatants of mitogen-activated EL4 thymoma cells. We used the proliferation of highly purified splenic B cells in the presence of anti-IgM as an assay in our purification procedure. The purification was also monitored for additional lymphokines using the factor-dependent cell lines CTLL-2 and FDC-P2. Interestingly, fractions containing BSF-1 always contained stimulatory activity for both IL-2- and IL-3-dependent cell lines, even though these fractions were devoid of IL-2 and IL-3 protein. N-terminal amino acid sequencing of homogeneous BSF-1 revealed a unique protein sequence completely different from that of murine IL-2 or IL-3. Based on these results, we conclude that BSF-1 is both a growth and differentiation factor that may have biologic effects beyond the B lymphocyte compartment.

### Materials and Methods

*Mice.* Female C57BL/6J mice were obtained from The Jackson Laboratory, Bar Harbor, ME, and were used at 8-12 wk of age.

<sup>1</sup> *Abbreviations used in this paper:* BSF-1, B cell-stimulating factor; DIEA-Ac, *N,N*-diisopropylethylamine; TFA, trifluoroacetic acid; TMS-silica, trimethylsilyl-silica.



**Preparation of B Lymphocytes.** Mice were pretreated with intraperitoneal injections of T24 rat anti-mouse Thy-1 mAb (12). 3 d later, the spleens were removed and treated in vitro with a cocktail of T24, GK1.5 rat anti-mouse L3T4 mAb, rabbit anti-mouse thymocyte serum (liver- and bone marrow-absorbed) and rabbit complement (Pel-Freeze Biologicals, Rogers, AR). The T cell-depleted spleen cells were then passed over Sephadex G10 to remove adherent cells, and the B lymphocytes were positively selected by panning on petri dishes coated with goat anti-mouse IgM (Cooper Biomedical, Inc., Malvern, PA). Purified B cells contained undetectable levels of T cells and >98% B cells as judged by flow cytometer analysis (Epics-C; Coulter Electronics, Inc., Hialeah, FL) using T24 and anti-L3T4 rat mAb, followed by FITC-conjugated rabbit anti-rat Ig (Becton Dickinson Immunocytometry Systems, Mountain View, CA) to detect T cells, and FITC-conjugated rabbit anti-mouse IgM to detect B cells. In addition, such a B cell population was found to be completely unresponsive to the T cell mitogen, Con A, but retained full responsiveness to the B cell mitogen LPS.

**Cell Lines.** EL4 thymoma cells were maintained in RPMI-1640 supplemented with 5% FCS, 50 U/ml penicillin, 50 µg/ml streptomycin and 2 mM glutamine. For the production of lymphokine-containing supernatants, EL4 cells were stimulated in serum-free medium with 1% PHA (PHA-M; Difco Laboratories, Detroit MI) and 10 ng/ml PMA (Sigma Chemical Co., St. Louis, MO). Cell-free supernatant was collected by centrifugation after 24 h.

CTLL-2 an IL-2-dependent murine T cell line (13) was maintained in Click's medium (Atlick Associates, River Falls, WI) containing 10% FCS, 50 U/ml penicillin, 50 µg/ml streptomycin, and 100 U/ml of IL-2 from a supernatant of rat spleen cells cultured 24 h with Con A (Pharmacia Fine Chemicals, Piscataway, NJ).

FDC-P2, a factor-dependent murine cell line derived from long-term bone marrow cultures (14) was maintained in RPMI-1640 containing 20% horse serum, 50 U/ml penicillin, 50 µg/ml streptomycin, 50 µM 2-ME, and 10% WEHI-3 cell line-conditioned medium.

32D, an IL-3-dependent murine hemopoietic cell line (15) was kindly provided by Dr. James Watson (Auckland Medical School, Auckland, New Zealand) and was maintained in Click's medium containing 10% WEHI-3 cell line-conditioned medium, 10% FCS, and antibiotics.

**Cellular Assays.** BSF-1 was assayed by its ability to stimulate the proliferation of purified B cells in the presence of a submitogenic concentration of goat anti-mouse IgM (7). B cells were cultured ( $10^5$  cells/culture) in 200 µl containing 3–5 µg/ml of affinity-purified goat anti-mouse IgM (Cooper Biomedical, Inc.) and serial dilutions of test sample. After 72 h, cultures received 2.0 µCi of [ $^3$ H]thymidine (75 Ci/mmol, New England Nuclear, Boston, MA) for 6 h, were harvested onto glass fiber filters, and incorporation of radioactivity was measured.

All proliferation assays using factor-dependent cell lines were performed using  $2 \times 10^5$  cells/culture in 100 µl of Click's medium containing 10% FCS, antibiotics, and test samples. After 24 h, cultures received 2.0 µCi of [ $^3$ H]thymidine (75 Ci/mmol) for 6 h, were harvested onto glass fiber filters, and incorporation of radioactivity was measured. For all proliferation assays, 1 U of activity was defined as the amount of lymphokine that induced 50% of maximal proliferation in 100 µl cultures. For example, if a sample induced 50% of maximal proliferation at a dilution of 1:20, then 1 U was said to be contained in one-twentieth of 100 µl, or 5 µl, and the sample said to contain 200 U/ml.

**Lymphokines.** A mouse IL-2 cDNA was cloned from a library prepared from LBRM-33 lymphoma cell mRNA by Immunex Corporation and was expressed in yeast using the yeast  $\alpha$  factor promoter and leader sequences to direct synthesis and secretion (16). rIL-2 was purified to homogeneity as previously described (17), and had a specific activity of  $10^9$  U/mg.

A mouse IL-3 cDNA was also cloned from the above-detailed library by Immunex Corporation, expressed in yeast, and resultant rIL-3 was purified to homogeneity (18). Purified murine rIL-3 had sp act of  $1.1 \times 10^{10}$  U/mg.

**BSF-1 Purification.** BSF-1 was purified from 71.5 liters of supernatant from PHA and

PMA-stimulated EL4 thymoma cells. BSF-1 activity was extracted from crude supernatants by adsorption to trimethylsilyl-silica (C-1, Sephadex International, Harbor City, CA) as previously described (19), with modifications. Crude supernatant was acidified with 0.1% trifluoroacetic acid (TFA), and trimethylsilyl (TMS)-silica added (10 g/liter). After stirring 1.5 h, the supernatant was decanted. The TMS-silica was poured into a column, washed with 20% acetonitrile with 0.1% TFA, and the BSF-1 was eluted with 75% acetonitrile and 0.1% TFA. Acetonitrile was removed by rotary evaporation, and the aqueous phase was adjusted to 5 mM sodium citrate, 50 mM NaCl, pH 5.5.

The TMS-silica-purified BSF-1 was fractionated on carboxymethyl cellulose (CM52; Whatman, Inc., Clifton, NJ). BSF-1 bound and was eluted with a linear NaCl gradient. The fractions containing BSF-1 were pooled, dialyzed against 20 mM Tris, pH 9.0, and fractionated on quaternary aminoethyl cellulose (QA52, Whatman, Inc.) to which BSF-1 bound, and was eluted with a linear NaCl gradient.

Reversed-phase HPLC fractionation of partially purified BSF-1 preparations was performed on a 4.6 × 250 mm Vydac 218TP (C<sub>18</sub>) column (The Separations Group, Hesperia, CA) using a Beckman Model 344 solvent delivery system. For purification step HPLC I (Table I), the column was equilibrated with 0.1% TFA in water at a flow rate of 0.8 ml/min. Fractions containing BSF-1 activity from the QA52 column were pooled, adjusted to pH 2 with TFA, and injected onto the column. The column was washed for 10 additional minutes with 0.1% TFA and then brought to 10% acetonitrile (containing 0.1% TFA) over 2 min. After an additional 8 min of equilibration at 10% acetonitrile, a linear gradient from 10 to 70% acetonitrile in 0.1% TFA was run over 60 min (1% per minute), and 1-min fractions were collected.

For HPLC II, active fractions from HPLC I were pooled and concentrated in vacuo to 250  $\mu$ l, and injected onto the same column, equilibrated with 50 mM acetic acid adjusted to pH 4.50 with *N,N*-diisopropylethylamine (DIEA-Ac), at a flow rate of 0.7 ml/min. After a 5 min wash with DIEA-Ac, the column was brought to 5% *n*-propanol over 2 min. After an additional 8 min of washing, a linear gradient from 5 to 40% *n*-propanol in DIEA-Ac over 70 min (0.5%/min) was run, and 1 min fractions were collected.

***N-terminal Protein Sequencing.*** Amino-terminal amino acid sequencing was performed on an Applied Biosystems Model 470A protein sequencer (Applied Biosystems Inc., Foster City, CA). Homogeneous BSF-1 from HPLC II fractions 61 and 62 (Fig. 1) were concentrated in vacuo to a final volume of 30  $\mu$ l and then spotted onto a conditioned sequencer filter. Sequencing and PTH amino acid analysis was performed as described previously (20).

***SDS-PAGE.*** Fractions from the purification steps were monitored by SDS-PAGE and subsequent silver staining, as described previously (21).

## Results

***Initial Purification.*** BSF-1 was purified from serum-free supernatants of PHA- and PMA-stimulated EL4 cells as described in Materials and Methods and Table I. BSF-1 was extracted from crude supernatants by adsorption to TMS-silica. BSF-1 was eluted in TFA/acetonitrile, and further fractionated by a combination of cation- (CM52) and anion- (QA52) exchange chromatography (Table I).

***HPLC.*** Biologically active fractions eluted from the QA52 column were pooled and fractionated by HPLC on a C<sub>18</sub> column as described in Materials and Methods. After one fractionation using a TFA/acetonitrile buffer system (HPLC I), an intense band at 18.4 kD was detected by silver staining on SDS-PAGE gels (Fig. 1, lane 1, SM), which correlated with BSF-1 biological activity. BSF-1 biological activity was eluted from the HPLC column in 42% acetonitrile.

The biologically active fractions from HPLC I were further fractionated using the same HPLC column equilibrated in a different buffer system (see Materials and Methods). As shown in Fig. 1, this fractionation yielded homogeneous BSF-

TABLE I  
*Purification of BSF-1*

Step	Total activity	Specific activity	Yield
	$U \times 10^{-6}$	$U/\mu g$	%
Crude	21.4	3	100
TMS-silica	19.7	100	92
CM52	13.7	2,500	64
QA52	5.6	26,000	26
HPLC I	5.1	127,100	24
HPLC II	1.6	328,000	8

BSF-1 was purified from 71.5 liters of crude EL4 supernatant as described in Materials and Methods. Units of BSF-1 activity were determined using purified B cells. Protein was measured using the Bradford protein assay (BioRad Laboratories, Richmond, CA).

1 protein as detected by silver-stained SDS-PAGE, and as confirmed by protein sequencing. The peak of BSF-1 biological activity eluted at 32% *n*-propanol. Although acceptable yields of protein were recovered at this step, some loss of biological activity was observed, which possibly was due to denaturation of the BSF-1 molecule in the propanol solvent. The observed sp act of this BSF-1 was  $3.28 \times 10^8$  U/mg.

**Protein Sequencing.** Fractions 61 and 62 (Fig. 1) from the HPLC II-purified preparation were subjected to amino-terminal protein sequencing. Only one sequence was obtained, consistent with the preparation of BSF-1 having been purified to homogeneity. The yields for each cycle of amino-terminal sequencing are shown in Table II. The initial yield, 56%, is based upon isoleucine in cycle two, due to typically poor yields for histidine residues (cycle one). The absolute yield of 67.4 pmol isoleucine in cycle two is consistent with the expected yield (50–70%) from the 120 pmol (2.2  $\mu g$ ) of BSF-1 applied to the sequencer filter. The first 20 residues, His-Ile-His-Gly-Cys-Asp-Lys-Asn-His-Leu-Arg-Glu-Ile-Ile-Gly-Ile-Leu-Asn-Glu-Val, were found to be a unique sequence when compared to previously published reports (22). The assignment of Cys to position 5 is tentative, since no signal was observed for cycle 5 and the protein was not modified before sequencing.

**Biological Activity.** All fractions throughout the purification were simultaneously monitored for BSF-1, IL-2 (CTLL proliferation), and IL-3 (FDC-P2 proliferation) activities. The majority of IL-2, as measured by CTLL proliferation, separated from BSF-1 on the CM52 column. However, subsequent to that step, all CTLL activity eluted precisely with the BSF-1 activity. EL4 supernatants contained no IL-3 protein, and all FDC-P2 activity eluted with BSF-1. The CTLL and FDC-P2 cell line response assays, using HPLC-purified BSF-1 fractions, are shown in Fig. 2. It is clear that purified BSF-1 has the capacity to score positively in both conventional IL-2 and IL-3 assays.

As seen in Fig. 3, several factor-dependent cell lines were compared for their responses to IL-2, IL-3, and BSF-1. CTLL, FDC-P2, and 32D all respond to BSF-1, but to a lesser extent (a lower maximum incorporation and a shallower slope) than to either IL-2 or IL-3.

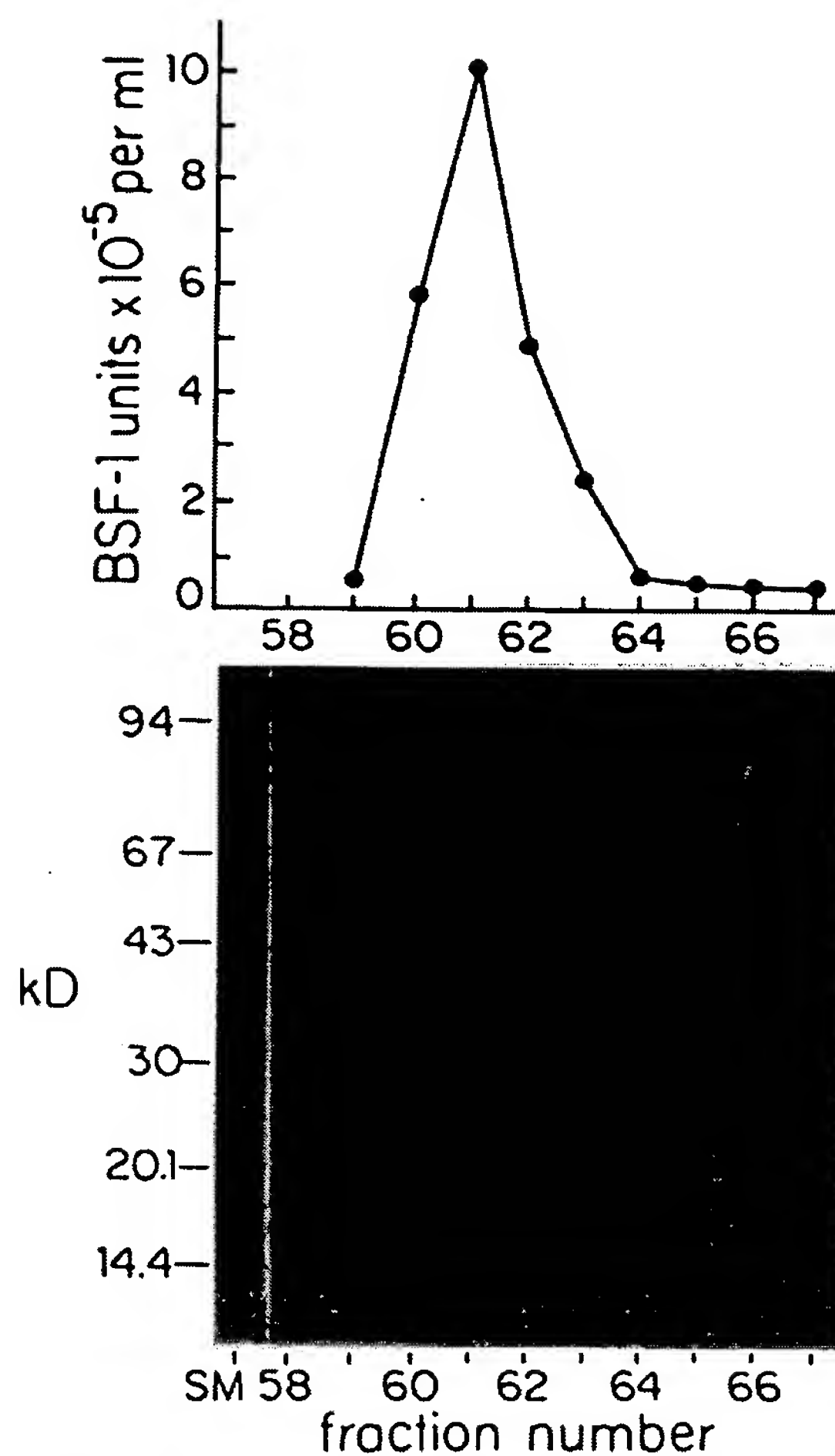


FIGURE 1. Purification of BSF-1 by HPLC. HPLC fractions from the final purification step (HPLC II) shown in Table I were assayed for BSF-1 activity (*top*) and analyzed by SDS-PAGE (*bottom*) as described in Materials and Methods. SM (starting material) in lane 1 is the pool of active fractions from HPLC I.

### Discussion

We have purified murine BSF-1 to homogeneity from supernatants of stimulated EL4 thymoma cells, and have determined its amino-terminal amino acid sequence. Homogeneous BSF-1 was found to have a molecular mass of 18.4 kD, with a specific activity of at least  $3.28 \times 10^8$  U/mg. In addition, this lymphokine was found to stimulate the proliferation of both IL-2- and IL-3-dependent cell lines.

BSF-1 was originally described as B cell growth factor because it stimulated partially purified splenic B lymphocytes to proliferate in the presence of submitogenic concentrations of goat anti-mouse IgM (7). However, recent evidence



TABLE II  
Protein Sequence from Purified Murine BSF-1

Amino acid	Yield (pmol) from cycle number:																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Asn	—*	—	—	—	—	—	—	45.5	—	—	—	—	—	—	—	—	—	23.5	—	—
Ser	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Thr	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Gln	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Gly	—	—	—	23.4	—	—	—	—	—	—	—	—	—	—	17.3	—	—	—	—	—
His	34.6	—	26.2	—	—	—	—	—	19.6	—	—	—	—	—	—	—	—	—	—	—
Ala	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Asp	—	—	—	—	—	36.6	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Arg	—	—	—	—	—	—	—	—	—	—	24.8	—	—	—	—	—	—	—	—	—
Glu	—	—	—	—	—	—	—	—	—	—	—	22.7	—	—	—	—	—	—	17.1	—
Tyr	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Val	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	26.6
Pro	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Met	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Ile	—	67.4	—	—	—	—	—	—	—	—	—	—	30.5	41.4	11.4	34.7	—	—	—	—
Leu	—	—	—	—	—	—	—	—	—	34.9	—	—	—	—	—	—	31.3	—	—	—
Phe	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Trp	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Lys	—	—	—	—	—	—	37.3	—	—	—	—	—	—	—	—	—	—	—	—	—
Cys	—	—	—	—	( <sup>‡</sup> )	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—

Sequence assigned to first 20 amino acids of BSF-1 is: His-Ile-His-Gly-Cys-Asp-Lys-Asn-His-Leu-Arg-Glu-Ile-Ile-Gly-Ile-Leu-Asn-Glu-Val.

\* Values <10 pmol are omitted for clarity.

<sup>‡</sup> Assigned as Cys due to lack of other signals; sulfhydryl groups were not modified before sequencing.

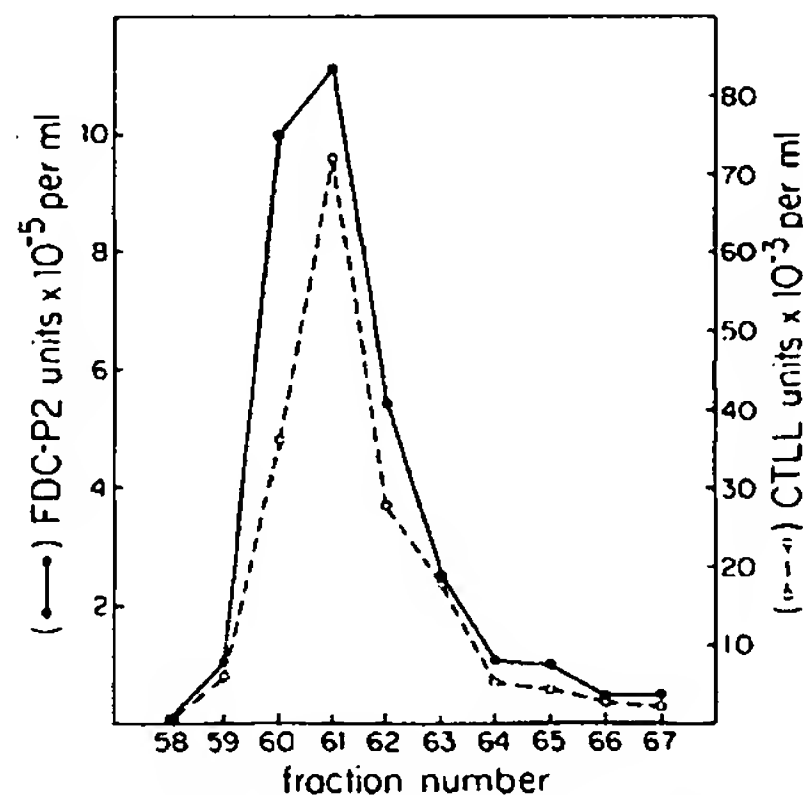


FIGURE 2. Assay of HPLC II fractions on CTLL and FDC-P2 cells. HPLC fractions from the final purification step (HPLC II) shown in Table I and Fig. 1 were assayed in the CTLL and FDC-P2 proliferation assays as described in Materials and Methods.

(10, 11) suggests that semipurified BSF-1 is not a direct growth factor for small resting B cells, but rather primes them for subsequent entry into S phase on encounter with anti-IgM, indicating that BSF-1 is a differentiation factor. Consistent with this role of BSF-1 are the findings that partially purified BSF-1 induces the cell surface expression of Ia on B cells (23, 24), and the secretion of IgG1 by LPS-activated B cells (25–27). Roehm et al. (24) have also observed that partially purified BSF-1 augments the antigen-presenting function of B cells.

Our observation that BSF-1 directly stimulates proliferation of IL-2- and IL-

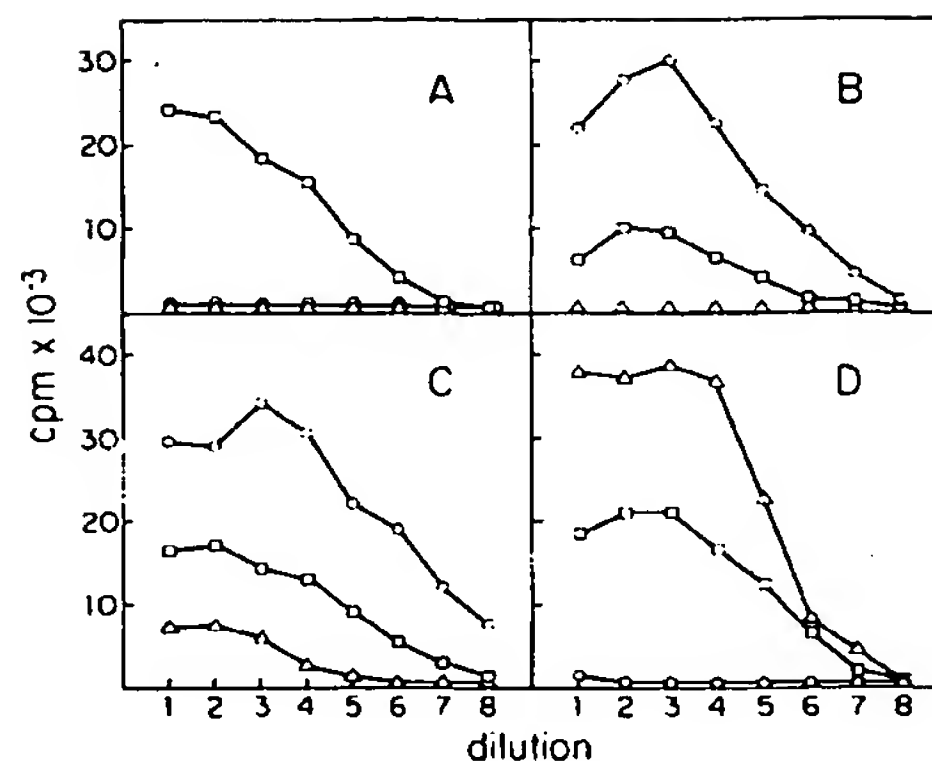


FIGURE 3. Response of B cells and factor-dependent cell lines to homogeneous lymphokines. Homogeneous BSF-1 ( $\square$ ), rIL-2 ( $\Delta$ ), and rIL-3 ( $\circ$ ) were tested for the stimulation of purified B cells (A), FDC-P2 (B), 32D (C), and CTLL (D) as described in Materials and Methods. Threefold dilutions of each lymphokine were tested starting at 500 U/ml and results reported in  $\text{cpm} \times 10^{-3}$ .

3-dependent cell lines indicates that BSF-1 is a growth factor as well as a differentiation factor. This should not be surprising, as similar results have been obtained in studies of other lymphokines. For example, IL-2 stimulates proliferation of T cells, however, it has recently become evident that IL-2 also elicits a variety of differentiated T cell functions (28–30). Moreover, granulocyte-macrophage-CSF (GM-CSF) has also been shown to activate differentiated macrophage (31) and granulocyte (32) effector functions, in addition to its traditional growth-stimulating activity on myeloid stem cells in the marrow.

Lymphokines that act on multiple lineages of cells are not without precedent. IL-2, once thought to act only on T lymphocytes, has recently (3, 4) been shown to drive proliferation of B cells as well. Additionally, several IL-3-dependent nonlymphoid cell lines have also been shown capable of responding to IL-2 (33, 34). Sanderson et al. (35) have reported that another B cell stimulating factor, BCGF-II, also acts as an eosinophil differentiation factor. One important caveat remains: although we have shown that BSF-1 stimulates non-B cell, factor-dependent cell lines to proliferate, it has yet to be determined whether this reflects a role of BSF-1 on normal cells, or whether those responses represent abnormal capabilities on the part of cells maintained in vitro for extended periods of time.

The ability of BSF-1 to act on IL-2- and IL-3-dependent cell lines suggests that BSF-1 has target cells beyond the B cell lineage, and may have significant functions other than the regulation of B cell growth and maturation.

### Summary

Murine B cell stimulating factor 1 (BSF-1) was purified to homogeneity from supernatants of a stimulated thymoma cell line. A protein of 18.4 kD with a unique N-terminal amino acid sequence was identified. BSF-1 had a sp act of at least  $3.28 \times 10^8$  U/mg. In addition to its B cell-stimulatory activity, BSF-1 also

stimulated the proliferation of several IL-2- and IL-3-dependent cell lines. We conclude that BSF-1 is both a growth factor and a differentiation factor. Finally, these results also suggest additional biologic properties of BSF-1 on lineages besides B lymphocytes.

We gratefully acknowledge the technical contributions of Carol Ramthun, Elizabeth Conrad, and Elizabeth McGrath, and we thank Judy Byce for typing the manuscript.

*Received for publication 18 February 1986.*

### References

1. Dutton, R. W., R. Falkoff, J. A. Hirst, M. Hoffman, J. W. Kappler, J. R. Kettman, J. F. Lesley, and D. Vann. 1971. Is there evidence for a non-antigen specific diffusible chemical mediator from the thymus-derived cell line in the initiation of the immune response? *In* Progress in Immunology. B. Amos, editor. Academic Press, New York. pp. 355-368.
2. Schimpl, A., and E. Wecker. 1972. Replacement to T cell function by a T cell product. *Nature (Lond.)* 237:15.
3. Mond, J. J., C. Thompson, F. D. Finkelman, J. Farrar, M. Schaefer, and R. J. Robb. 1985. Affinity-purified interleukin 2 induces proliferation of large but not small B cells. *Proc. Natl. Acad. Sci. USA* 82:1518.
4. Zubler, R. H., J. W. Lowenthal, F. Erard, N. Hashimoto, R. Devos, and H. R. MacDonald. 1984. Activated B cells express receptors for, and proliferate in response to, pure interleukin 2. *J. Exp. Med.* 160:1170.
5. Sidman, C. L., J. D. Marshall, L. D. Shultz, P. W. Gray, and H. M. Johnson. 1984.  $\gamma$ -interferon is one of several direct B cell-maturing lymphokines. *Nature (Lond.)* 309:801.
6. Leibson, H. J., M. Gester, A. Zlotnik, P. Marrack, and J. W. Kappler. 1984. Role of  $\gamma$ -interferon in antibody-producing responses. *Nature (Lond.)* 309:799.
7. Howard, M., J. Farrar, M. Hilfiker, B. Johnson, K. Takatsu, T. Hamaoka, and W. E. Paul. 1982. Identification of a T cell-derived B cell growth factor distinct from Interleukin 2. *J. Exp. Med.* 155:914.
8. Farrar, J. J., M. Howard, J. Fuller-Farrar, and W. E. Paul. 1983. Biochemical and physiochemical characterization of mouse B cell growth factor: A lymphokine distinct from interleukin 2. *J. Immunol.* 131:1838.
9. Dutton, R. W., G. D. Wetzel, and S. L. Swain. 1984. Partial purification and characterization of a BCGFII from EL4 culture supernatants. *J. Immunol.* 132:2451.
10. Rabin, E. M., J. Ohara, and W. E. Paul. 1985. B-cell stimulatory factor 1 activates resting B cells. *Proc. Natl. Acad. Sci. USA* 82:2935.
11. Oliver, K., R. J. Noelle, J. W. Uhr, P. H. Krammer, and E. S. Vitetta. 1985. B-cell growth factor (B-cell growth factor I or B-cell-stimulating factor, provisional I) is a differentiation factor for resting B cells and may not induce cell growth. *Proc. Natl. Acad. Sci. USA* 82:2465.
12. LeGros, G. S., R. L. Prestidge, and J. D. Watson. 1983. In vivo modulation of thymus-derived lymphocytes with monoclonal antibodies in mice. I. Effect of anti-Thy-1 antibody on the tissue distribution of lymphocytes. *Immunology* 50:537.
13. Gillis, S., and K. A. Smith. 1977. Long-term culture of tumor-specific cytotoxic T-cells. *Nature (Lond.)* 268:154.
14. Dexter, T. M., J. Garland, D. Scott, E. Scolnick, and D. Metcalf. 1980. Growth of factor-dependent hemopoietic precursor cell lines. *J. Exp. Med.* 152:1036.
15. Greenberger, J. S., R. J. Eckner, M. Sakakeeny, P. Marks, D. Reid, G. Nabel, A.

- Hapel, J. N. Ihle, and K. C. Humphries. 1983. Interleukin 3-dependent hemato-poietic progenitor cell lines. *Fed. Proc.* 42:2762.
16. Brake, A. J., J. P. Merryweather, D. G. Coit, U. A. Heberlein, F. R. Masiarz, G. T. Mullenbach, M. S. Urdea, P. Valenzuela, and P. J. Barr. 1984.  $\alpha$ -Factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA.* 81:4642.
  17. Stern, A. S., Y. C. E. Pan, D. L. Urdal, D. Y. Mochizuki, S. Dechiara, R. Blacher, J. Wideman, and S. Gillis. 1984. Purification to homogeneity and partial characterization of interleukin 2 from a human T-cell leukemia. *Proc. Natl. Acad. Sci. USA.* 81:871.
  18. Urdal, D. L., D. Mochizuki, P. J. Conlon, C. J. March, M. L. Remerowski, J. Eisenman, C. Ramthun, and S. Gillis. 1984. Lymphokine purification by reversed-phase high-performance liquid chromatography. *J. Chromatogr.* 296:171.
  19. Ohara, J., S. Lahet, J. Inman, and W. E. Paul. 1985. Partial purification of murine B cell stimulatory factor (BSF)-1. *J. Immunol.* 135:2518.
  20. March, C. J., and T. P. Hopp. 1986. In *Modern Methods in Protein Chemistry*. J. J. L'Italien, editor. Plenum Press, New York. In press.
  21. Kronheim, S., C. March, S. Erb, P. Conlon, D. Mochizuki, and T. P. Hopp. 1985. Human interleukin 1: Purification to homogeneity. *J. Exp. Med.* 161:490.
  22. Dayhoff, M. O., W. C. Barker, and L. T. Hunt. 1983. Establishing homologies in Protein Sequences. *Methods Enzymol.* 91:524.
  23. Noelle, R., P. H. Krammer, J. Ohara, J. W. Uhr, and E. S. Vitetta. 1984. Increased expression of Ia antigens on resting B cells: An additional role for B-cell growth factor. *Proc. Natl. Acad. Sci. USA.* 81:6149.
  24. Roehm, N. W., H. J. Leibson, A. Zlotnik, J. Kappler, P. Marrack, and J. C. Cambier. 1984. Interleukin-induced increase in Ia expression by normal mouse B cells. *J. Exp. Med.* 160:679.
  25. Vitetta, E. S., J. Ohara, C. D. Myers, J. E. Layton, P. H. Krammer, and W. E. Paul. 1985. Serological, biochemical, and functional identity of B cell-stimulatory factor 1 and B cell differentiation factor for IgG1. *J. Exp. Med.* 162:1726.
  26. Yuan, D., E. A. Weiss, J. E. Layton, P. H. Krammer, and E. S. Vitetta. 1985. Activation of the  $\gamma 1$  gene by lipopolysaccharide and T cell-derived lymphokines containing a B cell differentiation factor for IgG1 (BCDF $\gamma$ ). *J. Immunol.* 135:1465.
  27. Jones, S., J. Layton, P. H. Krammer, E. S. Vitetta, and P. W. Tucker. 1985. The effect of T cell-derived lymphokines on the levels of isotype-specific RNA in normal B cells. *J. Mol. Cell. Immunol.* 2:143.
  28. Reem, G. H., and N.-H. Yeh. 1984. Interleukin 2 regulates expression of its receptor and synthesis of gamma interferon by human T lymphocytes. *Science (Wash. DC).* 225:429.
  29. Farrar, W. L., H. M. Johnson, and J. J. Farrar. 1981. Regulation of the production of immune interferon and cytotoxic T lymphocytes by interleukin 2. *J. Immunol.* 126:1120.
  30. Welte, K., M. Andreeff, E. Platzer, K. Holloway, B. Y. Rubin, M. A. S. Moore, and R. Mertelsmann. 1984. Interleukin 2 regulates the expression of Tac antigen on peripheral blood T lymphocytes. *J. Exp. Med.* 160:1390.
  31. Grabstein, K. H., D. L. Urdal, R. J. Tushinski, D. Y. Mochizuki, V. L. Price, M. A. Cantrell, S. Gillis, and P. J. Conlon. 1986. Induction of macrophage tumoricidal activity by granulocyte-macrophage colony stimulating factor. *Science (Wash. DC).* 232:506.
  32. Lopez, A. F., N. A. Nicola, A. W. Burgess, D. Metcalf, F. L. Battye, W. A. Sewell,



- and M. Vadas. 1983. Activation of granulocyte cytotoxic function by purified mouse colony-stimulating factors. *J. Immunol.* 131:2983.
33. Koyasu, S., J. Yodoi, T. Nikaido, Y. Tagaya, Y. Taniguchi, T. Honjo, and I. Yahara. 1986. Expression of interleukin 2 receptors on interleukin 3-dependent cell lines. *J. Immunol.* 136:984.
34. Le Gros, G. S., S. Gillis, and J. D. Watson. 1985. Induction of IL2 responsiveness in a murine IL-3-dependent cell line. *J. Immunol.* 135:4009.
35. Sanderson, C. J., A. O'Garra, D. J. Warren, and G. G. B. Klaus. 1986. Eosinophil differentiation factor also has B-cell growth factor activity: proposed name interleukin 4. *Proc. Natl. Acad. Sci. USA.* 83:437.

## Homodimerization of Interleukin-4 Receptor $\alpha$ Chain Can Induce Intracellular Signaling\*

(Received for publication, April 26, 1996, and in revised form, August 5, 1996)

Winfried Kammer†, Antje Lischke‡, Richard Moriggl§, Bernd Groner§, Andrew Ziemiecki¶, Christine B. Gurniak||, Leslie J. Berg||, and Karlheinz Friedrich‡\*\*

From the ‡Theodor-Boveri-Institut für Biowissenschaften (Biozentrum), Physiologische Chemie II, Am Hubland, D-97074 Würzburg, Federal Republic of Germany, the §Institute for Experimental Cancer Research, Tumor Biology Center, Breisacher Strasse 117, D-79106 Freiburg, Federal Republic of Germany, the ¶Laboratory for Clinical & Experimental Cancer Research, Tiefenaustasse 120, CH-3004 Bern, Switzerland, and the ||Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138

The possible role of homodimerization events in intracellular signal transduction triggered by the bipartite human interleukin-4 receptor was addressed. We generated cell lines functionally expressing derivatives of the two receptor subunits  $\alpha$  and  $\gamma$ , which allow for a specific and background-free experimental induction of intracellular homo- and heterodimers. A heterodimer of  $\alpha$  and  $\gamma$  released an intracellular signal, whereas a  $\gamma$ - $\gamma$  homodimer did not. Unexpectedly, we found the intracellular domain of interleukin-4 receptor  $\alpha$  chain to evoke cell proliferation and activation of tyrosine kinase Jak1 as well as of transcription factor Stat6 upon homodimerization. Both recruitment of the common  $\gamma$  chain and activation of kinase Jak3 were shown to be dispensible for these processes.

Interleukin-4 (IL-4)<sup>1</sup> is a pleiotropic immune regulator with a pivotal role in certain allergic processes (1). The bipartite IL-4 receptor comprises the interleukin-4 receptor  $\alpha$  chain (IL-4R $\alpha$ ) (2) and the common  $\gamma$  receptor chain ( $\gamma$ c) (3, 4). Both receptor subunits belong to the cytokine receptor superfamily (5) and are shared by other cytokines;  $\gamma$ c is also part of the receptors for IL-2, IL-7, IL-9, and IL-15 (6), and IL-4R $\alpha$  contributes to the IL-13 receptor (7, 8).

Ligand-induced juxtaposition of the cytoplasmic domains of IL-4R $\alpha$  and  $\gamma$ c is believed to be a mandatory step in intracellular signaling which involves recruitment and activation of

kinases Jak1 and Jak3 (9, 10), transcription factor Stat6 (11), and the adaptor molecule IRS-2 (12). However, the architecture of the IL-4R complex as well as the molecular mechanisms underlying the specificity of IL-4-induced signal transduction are to date poorly understood.

Making use of the strictly species-specific interaction of interleukin-4 with IL-4R $\alpha$  chain, factor-dependent murine cells were rendered responsive to hIL-4 by expressing human IL-4R $\alpha$  (2, 13–16). An implication of these results is the ability of human IL-4 to activate IL-4 receptor complexes containing either human or murine common  $\gamma$  chain, thus complicating an analysis of the composition of the signaling competent receptor subunit assembly.

In order to study the role of receptor chain dimerization events in signal release by the hIL-4R complex, we generated an expression system for receptor subunits that allowed us to experimentally induce specific and background-free intracellular hetero- and homodimerization.

Our results show that the juxtaposition of two intracellular domains of IL-4R $\alpha$  can act as the trigger of specific signaling, including the activation of Jak1 and Stat6 and the induction of cell proliferation. Surprisingly, a hitherto assumed participation of the cytoplasmic portion of common  $\gamma$  chain and of  $\gamma$ c-associated kinase Jak3 is not required.

### MATERIALS AND METHODS

**DNA Manipulations, Stable Transfection of Mouse Cells, and Detection and Quantification of Receptor Expression**—Recombinant DNA work was performed according to standard procedures (17). The murine pre-B cell line Ba/F3 (18) has been described. BAF-4 $\alpha$ -py, a Ba/F3 derivative expressing both subunits of the human IL-4R, is identical to BAF-4R $\gamma$  (16).

Hybrid receptor genes were generated by polymerase chain reaction amplification of gene fragments from pKCR-py (16) encoding the epitope-tagged extracellular domain and transmembrane/intracellular domain of human  $\gamma$ c and exchanging them for the corresponding fragments (*Bam*HI/*Xho*I or *Xho*I/*Hind*III) in pKCR-4 $\alpha$ . The resulting expression plasmids pKCR-4 $\alpha$ / $\gamma$  and pKCR-py/ $\alpha$  were cotransfected into Ba/F3 cells as described (16).

Surface expression of receptor constructs was assayed by reacting intact cells with antibodies X 14/38 (16, 19) or P5D4 (20) specific for the extracellular portions of recombinant hIL-4R $\alpha$  or epitope-tagged human  $\gamma$ c, respectively, and subsequent detection of bound antibodies by peroxidase-coupled secondary antibodies as detailed elsewhere (21). Briefly, 10<sup>5</sup> cells in a microtiter well were incubated on ice for 30 min with 5  $\mu$ g of antibody in a volume of 50  $\mu$ l of phosphate-buffered saline/3% bovine serum albumin. After washing twice, cells were resuspended in 100  $\mu$ l of a 100  $\mu$ g/ml solution of peroxidase-conjugated goat anti-mouse IgG (Dianova) and kept on ice for 30 min. Cell-bound secondary antibody was detected by transferring the cells to 50  $\mu$ l of a solution containing 0.1 M Tris/HCl, pH 8.5, 2.5 mM 3-aminophthalhydrazide (Fluka), 400  $\mu$ M *p*-coumaric acid (Sigma), 5.4 mM H<sub>2</sub>O<sub>2</sub> and measuring elicited chemiluminescence using a MicroLumat LB 96P. Quantitation of surface-bound antibody molecules was achieved by relating the determined intensity of luminescence to a calibration series of samples containing known concentrations of peroxidase.

**Cell Culture, Cytokines, and Proliferation Assay**—Cell maintenance and preparation of hIL-4 and mutant Y124D has been described previously (16). Recombinant murine IL-4 was purchased from Sigma. Cytokine-induced proliferation of cell lines was measured by [<sup>3</sup>H] thymidine incorporation into *de novo* synthesized DNA as described (16).

**Immunoprecipitation, Immunoblotting, and Chemical Cross-linking**—Samples of 3  $\times$  10<sup>7</sup> cells were incubated at 37 °C for 10 min in 1 ml of RPMI containing no cytokine, 7 nM of IL-4, or 50 nM of antibody P5D4 and subsequently lysed as described (16). Cleared lysates were incubated with 1–5  $\mu$ g of specific antibody. Antibodies used for immunoprecipitations were 4G10 (anti-phosphotyrosine, Upstate Biotechnology),

\* This work was supported by Deutsche Forschungsgemeinschaft through SFB 176. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

\*\* To whom correspondence should be addressed. Tel.: 49-931-888-4124; Fax: 49-931-888-4113.

<sup>1</sup> The abbreviations used are: IL, interleukin; hIL-4, human interleukin-4; hIL-4R, human interleukin-4 receptor; hIL-4R $\alpha$ , human interleukin-4 receptor  $\alpha$  chain;  $\gamma$ c, common receptor  $\gamma$  chain; Jak, janus kinase; Stat, signal transducer and activator of transcription;  $\alpha$ -, anti-

E34-1 (22) and anti-Jak1 rabbit serum (23). Immunocomplexes were precipitated from lysates with 50  $\mu$ l of anti-mouse IgG-agarose or protein A-Sepharose (Sigma) and assayed as described (16) using peroxidase-conjugated antibody RC20 (Transduction Laboratories) at a final concentration of 0.1  $\mu$ g/ml. Iodination of hIL-4, cross-linking of radioligand to cell-surface receptors, and analysis of immunoprecipitated complexes by electrophoresis was carried out as described (19).

**Analysis of Stat Activation by Electrophoretic Mobility Shift Assay—**Whole cell extracts were prepared from cells stimulated with IL-4 or antibody as described above by suspension of cell pellets in a buffer containing 20 mM Hepes, pH 7.9, 400 mM NaCl, 1 mM EDTA, 20% glycerol, 1 mM dithiothreitol, 0.2 mM phenylmethylsulfonyl fluoride, 5  $\mu$ g/ml leupeptin, 5  $\mu$ g/ml, and 100  $\mu$ M sodium *ortho*-vanadate followed by three freeze-thaw cycles and centrifugation at 4 °C and 14,000 rpm for 15 min. Supernatants equivalent to  $10^5$  cells were used for bandshift assays performed as described (24). As a probe, the Stat6-binding sequence 5'-GTCAACTTCCCAAGAACAGAA-3' derived from the human  $\gamma$ -promoter (25) end-labeled with polynucleotide kinase to a specific activity of 8,000 cpm/fmol was applied. Supershift of Stat6 containing complexes was achieved by adding to the binding reactions before electrophoretic mobility shift assay 1  $\mu$ g of a chicken antibody directed to amino acids 637–847 of murine Stat6.<sup>2</sup>

## RESULTS AND DISCUSSION

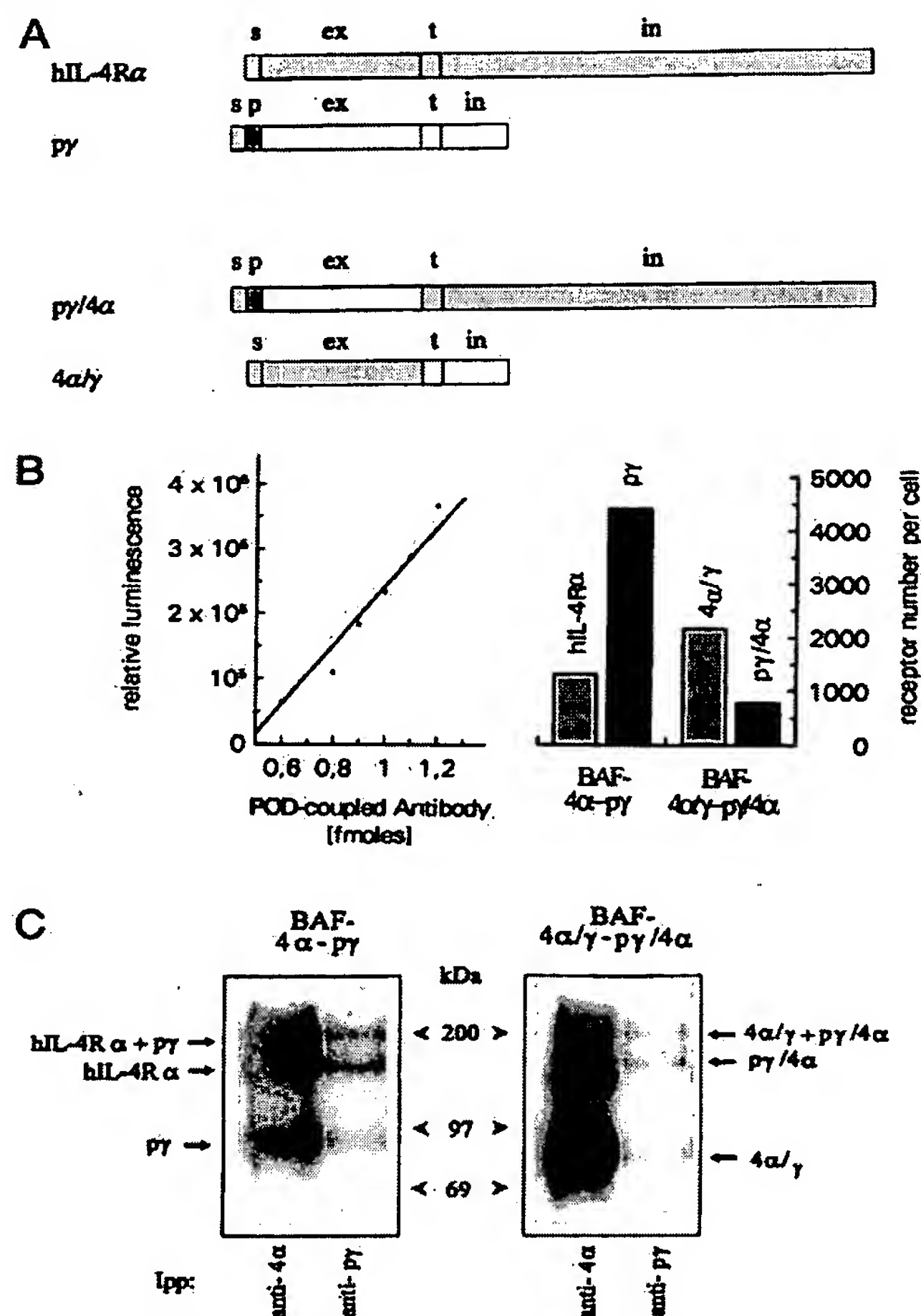
We intended to reconstitute in murine cells a functional interleukin-4 receptor complex activable exclusively by human IL-4, which would not evoke any background signaling due to interference with the endogenous murine IL-4 receptor. To this end, we generated a pair of expression constructs encoding hybrid receptor chains derived from hIL-4R $\alpha$  and  $\gamma$  with mutually exchanged intracellular domains (Fig. 1A) and introduced it into the murine pre-B cell line Ba/F3.

One clone expressing both 4 $\alpha/\gamma$  and  $\gamma$ /4 $\alpha$  chimeras was termed BAF-4 $\alpha/\gamma$ - $\gamma$ /4 $\alpha$ . The number of surface-expressed receptor molecules per cell was determined in comparison with cell line BAF-4 $\alpha$ - $\gamma$  bearing both subunits of the authentic human IL-4R (Fig. 1B). As measured by the binding of specific antibodies recognizing the extracellular receptor domains, in both cell lines surface expression of the receptor chain comprising the intracellular domain of  $\gamma$  was considerably higher than that of the subunit bearing the intracellular part of hIL-4R $\alpha$ . Irrespective of the “authentic” or “cross-over” composition of the heterologous subunits, similar hIL-4 binding receptor complexes could be formed in both cell lines as revealed by immunoprecipitation of receptor chains cross-linked to radiolabeled hIL-4 (Fig. 1C).

To test if the bipartite human IL-4R with exchanged cytoplasmic domains was capable of transmitting specific signals to the cell interior, we measured IL-4-induced cell proliferation. When stimulated with hIL-4, BAF-4 $\alpha/\gamma$ - $\gamma$ /4 $\alpha$  cells expressing the combination of hybrid receptors, like BAF-4 $\alpha$ - $\gamma$  cells, showed a proliferative response (Fig. 2A).

In BAF-4 $\alpha$ - $\gamma$  cells, hIL-4 mutant Y124D evoked 60% of the DNA synthesis induced by wild type IL-4. We have previously shown that this degree of reactivity is due to preferential interaction of Y124D with murine  $\gamma$  (16). When assaying BAF-4 $\alpha/\gamma$ - $\gamma$ /4 $\alpha$  cells, we found, as earlier observed with human IL-4 reactive cells (19), only 30% of wild type activity for hIL-4 variant Y124D. This result indicates that hIL-4 cross-over receptor, as anticipated and unlike its authentic counterpart, precludes the formation of productive receptor complexes involving endogenous murine common  $\gamma$  chain.

Stimulation with hIL-4 resulted in equivalent patterns of tyrosine-phosphorylated proteins in the two cell lines (Fig. 2B). The intracellular domain of hIL-4R $\alpha$  is a major substrate of ligand-induced phosphorylation as revealed by specific immunoprecipitation (data not shown). Moreover, the modified hIL-4 receptor was found to recapitulate hIL-4-specific activation of



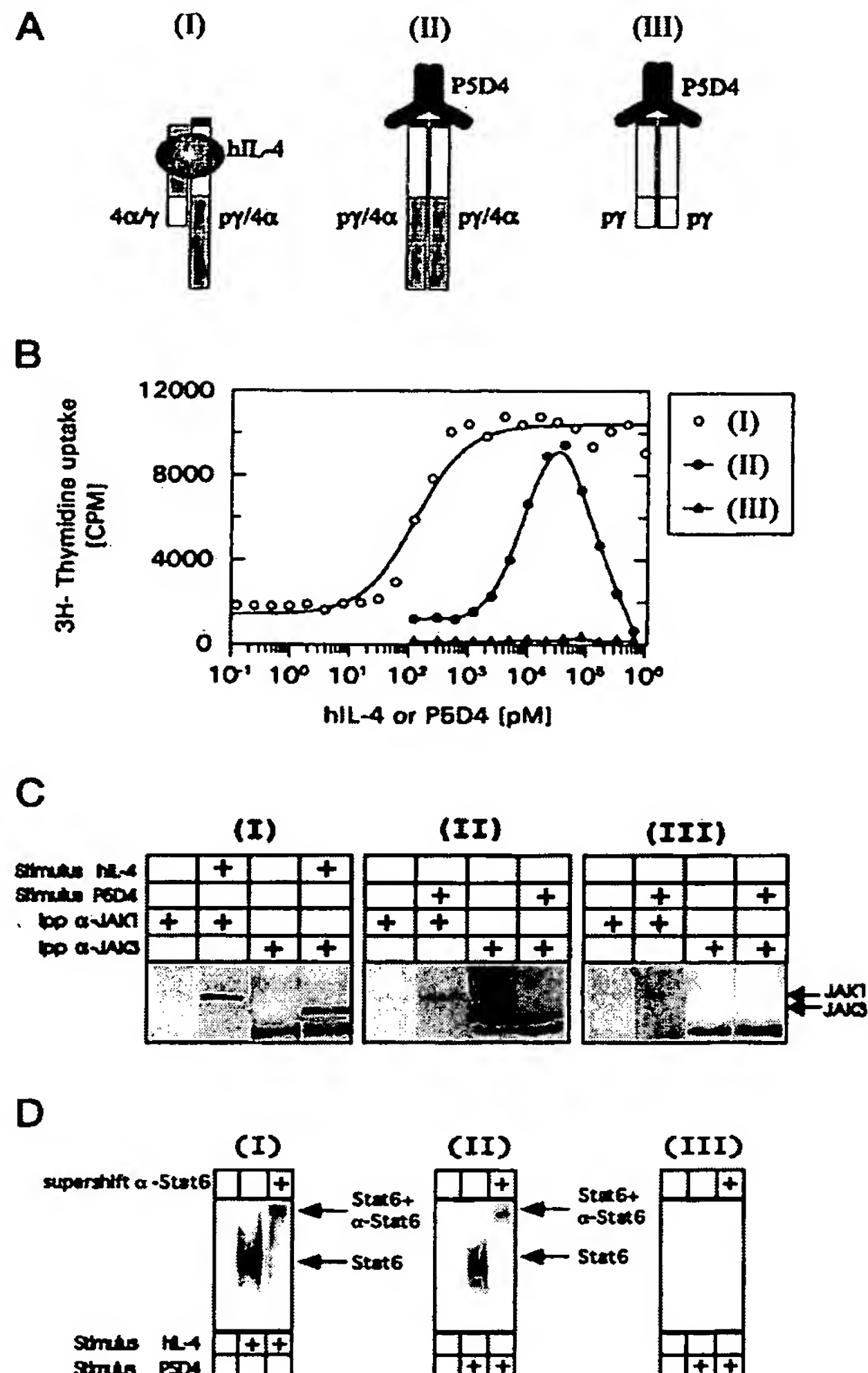
**FIG. 1. Generation and characterization of cell lines expressing hIL-4 receptor constructs.** A, schematic representation of bipartite authentic hIL-4 receptor expressed in BAF-4 $\alpha$ -py cell (*top pair*) in hIL-4 cross-over receptor expressed in BAF-4 $\alpha/\gamma$ -py/4 $\alpha$  cells (*bottom pair*). s, signal peptide; ex, extracellular domain; t, transmembrane domain; in, intracellular domain; p, epitope tag recognized by antibody P5D4. B, analysis of receptor chain surface expression in BAF-4 $\alpha$ -py and BAF-4 $\alpha/\gamma$ -py/4 $\alpha$  cells. Samples of  $10^5$  cells were reacted with antibodies directed to the extracellular domains of hIL-4R $\alpha$  or epitope-tagged  $\gamma$ c, respectively, washed, and stained with peroxidase-conjugated secondary antibody as described under “Materials and Methods.” Numbers of bound enzyme molecules per cell equivalent to receptor chain copies were determined by quantification of elicited chemiluminescence and correlation of the signal intensity with a calibration series obtained by measuring luminescence produced by different known amounts of peroxidase under assay conditions. C, analysis of ligand-receptor complexes formed on BAF-4 $\alpha$ -py and BAF-4 $\alpha/\gamma$ -py/4 $\alpha$  cells. After chemical cross-linking of  $^{125}$ I-hIL-4 to the two cell lines, receptor complexes were immunoprecipitated using the indicated antibodies and subsequently resolved and visualized by SDS-polyacrylamide gel electrophoresis and autoradiography. Radioligand cross-linked receptor chains and complexes are marked with arrows.

janus kinases Jak1 and Jak3.

We next employed the model receptor system to address the individual roles of the IL-4 receptor subunits in signaling. The ligand and antibody binding properties of the functionally expressed receptor constructs enabled us to specifically induce all three possible intracellular receptor dimers (Fig. 3A). In BAF-4 $\alpha/\gamma$ -py/4 $\alpha$  cells, not only hIL-4-induced heterodimerization of the two intracellular receptor domains but surprisingly also antibody-mediated cytoplasmic homodimerization of hIL-4R $\alpha$  via the extracellular P5D4 epitope tag lead to cell proliferation

<sup>2</sup> M. Heim and R. Moriggl, unpublished data.





**FIG. 3. Directed cytoplasmic hetero- and homodimerization of hIL-4R subunits and resulting signal transduction.** *A*, schematic representation of stimulus-induced dimerization events. *I*, hIL-4-induced selective intracellular heterodimerization of hIL-4R $\alpha$  and human  $\gamma$ c in BAF-4a/ $\gamma$ -p $\gamma$ /4a cells. *II*, intracellular homodimerization of hIL-4R $\alpha$  via extracellular antibody and epitope tag in BAF-4a/ $\gamma$ -p $\gamma$ /4a cells. *III*, antibody-induced homodimerization of  $\gamma$ c in BAF-4a-p $\gamma$  cells. *B*, cell proliferation evoked by the stimuli depicted under *A*. The respective cell lines were incubated with the indicated concentrations of hIL-4 (*I*) or antibody P5D4 (*II* and *III*) for 24 h before [ $^3$ H]thymidine uptake was measured. *C* and *D*, activation of Jak kinases (*C*) and activation of Stat6 (*D*) by the stimuli depicted under *A*. The respective cell lines were stimulated for 10 min with 10 nM hIL-4 (*I*) or 100 nM P5D4 (*II* and *III*). Cells were then lysed and subjected to immunoprecipitations with anti-Jak antibodies and probing with anti-phosphotyrosine antibody (*C*) or to a band shift assay using a labeled probe derived from the I $\epsilon$ -promoter (*D*) as described under "Material and Methods."

Comparing the activation of janus kinases known to be involved in IL-4 receptor complex function by hetero- or homodimerization, respectively (Fig. 3C), we found that antibody-induced intracellular homodimerization of hIL-4R $\alpha$  results in tyrosine phosphorylation of Jak1 but not of Jak3. Homodimerization of intracellular  $\gamma$ c does not lead to a detectable phosphorylation of Jak1 or Jak3, whereas the heterodimer of  $\alpha$  and  $\gamma$  evokes the activation of both kinases. Activation of Stat6, as assayed by its property to bind to a cognate DNA sequence derived from the I $\epsilon$ -promoter, is induced not only by an intracellular heterodimer of  $\alpha$  and  $\gamma$  but also by an  $\alpha$ - $\alpha$  homodimer (Fig. 3D). From these results we conclude that both the cytoplasmic domain of  $\gamma$ c and activated Jak3 are not mandatory for Stat activation and for the onset of a signaling cascade leading to cell proliferation. The essential trigger for the release of these events is rather the juxtaposition of two intracellular domains of IL-4 receptor  $\alpha$  chain and the concomitant activation of Jak1 by tyrosine phosphorylation.

tor complexes and how specificity of these receptors is achieved despite their sharing of a subunit. The only defined biochemical function of  $\gamma_c$  is the recruitment of Jak3 to the receptor complex (27). This very process, however, has been shown not to be essential for the specific activity of the IL-2 receptor; it can rather be replaced by artificially introducing Jak2 into the assembly (28). In this report we show that in the human IL-4 receptor system neither  $\gamma_c$  and Jak3 nor any substitute is required for the release of an intracellular signal if the intracellular domain of hIL-4R $\alpha$  is experimentally homodimerized.

Our results raise new questions about the role of  $\gamma c$  in the



function of the IL-4 receptor. The interaction of ligand with the extracellular domains of both IL-4R $\alpha$  and  $\gamma$ c is necessary for IL-4-induced signal transduction, because mutant forms of hIL-4 defective in contacting  $\gamma$ c fail to stimulate cell proliferation (26, 29). Functional properties of the intracellular domain of  $\gamma$ c in the activation of this particular receptor system have not yet been addressed. Our data indicate that it is not involved in the release of intracellular signals specific for IL-4 and support the notion of a more general role for  $\gamma$ c in the formation of the signaling competent IL-4R and probably also other cytokine receptor complexes. In ligand-induced IL-4R activation, one function of  $\gamma$ c and Jak3 could be the promotion of a transient assembly of two or more copies of hIL-4R $\alpha$ , a situation which in turn would lead to specific intracellular signal transduction. Alternatively, in the natural receptor complex,  $\gamma$ c-mediated recruitment of Jak3 might result in an activation of Jak1, an event that in our model experiment is mimicked by the juxtaposition of two Jak1 molecules and serves as the master trigger for the various activities of hIL-4R $\alpha$ . A more general version of such an interpretation of exchangeable Jaks in the hIL-4R complex would be the view that ligand-induced intracellular apposition of several combinations of two Jak molecules would suffice to evoke cell proliferation and the other reactions observed. In this scenario, the major function of the specific receptor chain (here: hIL-4R $\alpha$ ) would be to provide recognition sites for Stats and other downstream components that upon Jak-driven activation mediate the particular effects of IL-4. Directed homodimerization of  $\gamma$ c does not result in similar activities because of its lack of recognition sites for downstream signaling molecules. Also in line with such an explanation would be the notion of cytokine receptor signal transduction being relatively unselective and flexible in terms of interactions between receptor chains and intracellular binding partners. This would imply that the main event regulating specificity in cytokine signaling is the recognition between receptor and ligand and the thereby cross-linked combination of receptor subunits.

To discriminate between the two principal explanations compatible with our results (involvement of receptor multimers in "natural" hIL-4 receptor activation or low specificity of Jak activity combined with recruitment of signaling molecules by hIL-4R $\alpha$  via specific recognition sites), careful investigation of the stoichiometric subunit composition of the active hIL-4 receptor complex and a mutational analysis of the cytoplasmic portion of  $\gamma$ c in this context are necessary. Also, the molecular

details of Jak recognition, activation, and specificity in the hIL-4R assembly have to be addressed.

**Acknowledgments**—The expert technical assistance of C. Müller is gratefully acknowledged. We thank W. Sebald for generous support and cytokines and T. Kreis for antibody P5D4.

## REFERENCES

1. Paul, W. E. (1991) *Blood* **77**, 1859–1870
2. Idzerda, R. L., March, C. J., Mosley, B., Lyman, S. D., Bos, T. V., Gimpel, S. D., Din, W. S., Grabstein, K. H., Widmer, M. B., Park, L. S., Cosman, D., and Beckman, M. P. (1990) *J. Exp. Med.* **171**, 861–973
3. Russell, S. M., Keegan, A. D., Harada, N., Nakamura, Y., Noguchi, M., Leland, P., Friedman, M. C., Miyajima, A., Puri, R. K., Paul, W. E., and Leonard, W. J. (1993) *Science* **262**, 1880–1883
4. Kondo, M., Takeshita, T., Ishii, N., Nakamura, M., Watanabe, S., Arai, K., and Sugamura, K. (1993) *Science* **262**, 1874–1879
5. Bazan, J. F. (1990) *Immunol. Today* **11**, 350–354
6. Sugamura, K., Asao, H., Kondo, M., Tanaka, N., Ishii, N., Nakamura, M., and Takeshita, T. (1995) *Adv. Immunol.* **59**, 225–277
7. Smerz-Bertling, C., and Duschl, A. (1995) *J. Biol. Chem.* **270**, 966–970
8. Obiri, N. I., Debinski, W., Leonard, W. J., and Puri, R. K. (1995) *J. Biol. Chem.* **270**, 8797–8804
9. Yin, T., Tsang, M. L.-S., and Yang, Y.-C. (1994) *J. Biol. Chem.* **269**, 26614–26617
10. Witthuhn, B. A., Silvennoinen, O., Miura, O., Lai, K. S., Cwik, C., Liu, E. T., and Ihle, J. N. (1994) *Nature* **370**, 153–157
11. Hou, J., Schindler, U., Henzel, W. J., Ho, T. C., Brasseur, M., and McKnight, S. L. (1994) *Science* **265**, 1701–1706
12. Keegan, A. D., Nelms, K., White, M., Wang, L.-M., Pierce, J. H., and Paul, W. E. (1994) *Cell* **76**, 811–820
13. Harada, N., Yang, G., Miyajima, A., and Howard, M. (1992) *J. Biol. Chem.* **267**, 22752–22758
14. Seldin, D. C., and Leder, P. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 2140–2144
15. Koettwitz, K., and Kalthoff, F. S. (1993) *Eur. J. Immunol.* **23**, 988–991
16. Lischke, A., Kammer, W., and Friedrich, K. (1995) *Eur. J. Biochem.* **234**, 100–107
17. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
18. Palacios, R., and Steinmetz, M. (1985) *Cell* **41**, 727–734
19. Bönsch, D., Kammer, W., Lischke, A., and Friedrich, K. (1995) *J. Biol. Chem.* **270**, 8452–8457
20. Kreis, T. E., and Lodish, H. F. (1986) *Cell* **46**, 929–937
21. Lischke, A., Pagany, M., Kammer, W., and Friedrich, K. (1996) *Anal. Biochem.* **236**, 322–326
22. Gurniak, C. B., and Berg, L. J. (1996) *Blood* **87**, 3151–3160
23. Wilks, A. F., Harpur, A. G., Kurban, R. R., Ralph, S. J., Zürcher, G., and Ziemiecki, A. (1991) *Mol. Cell. Biol.* **11**, 2057–2065
24. Wakao, H., Schmitt-Ney, and Groner, B. (1992) *J. Biol. Chem.* **267**, 16365–16370
25. Köhler, I., and Rieber, E. P. (1993) *Eur. J. Immunol.* **23**, 3066–3071
26. Kruse, N., Shen, B.-J., Arnold, S., Tony, H.-P., Müller, T., and Sebald, W. (1993) *EMBO J.* **13**, 5121–5129
27. Nelson, B. H., Lord, J. D., and Greenberg, P. D. (1996) *Mol. Cell. Biol.* **16**, 309–317
28. Lai, S. Y., Xu, W., Gaffen, S. L., Liu, K. D., Longmore, G. D., Greene, W. C., and Goldsmith, M. A. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 231–235
29. Duschl, A. (1995) *Eur. J. Biochem.* **228**, 305–310

**INTERLEUKIN 5 ENHANCES INTERLEUKIN 2-MEDIATED  
LYMPHOKINE-ACTIVATED KILLER ACTIVITY**

By TOMOKAZU AOKI, HARUHIKO KIKUCHI, SHIN-ICHI MIYATAKE,  
YOSHIFUMI ODA, KOICHI IWASAKI, TOSHIKI YAMASAKI,  
TATSUO KINASHI,\* AND TASUKU HONJO\*

*From the Department of Neurosurgery and \*Medical Chemistry, Faculty of Medicine,  
Kyoto University, Sakyo-ku, Kyoto 606, Japan*

After 3–6 d of in vitro incubation with IL-2, human PBL and mouse splenic lymphocytes become cytotoxic to a variety of tumor cells, including not only NK-sensitive, but also NK-resistant tumor cells (1–3). These IL-2-dependent lymphocytes consist of a heterogeneous population of effectors (4–6). They also show non-MHC-restricted cytotoxicity, which has been referred to as the lymphokine-activated killer (LAK) phenomenon (2).

On the other hand, murine IL-5 (mIL-5) was previously described as a T cell-replacing factor (TRF) (7, 8) and as a B cell growth factor II (9) based on its induction of proliferation and Ig secretion by activated B cells. This lymphokine has been shown to exert several other effects, including the induction of growth and differentiation of eosinophils from bone marrow stem cells (10), the induction of differentiation of peanut agglutinin-positive (PNA<sup>+</sup>) thymocytes into CTL (11), and the induction of mIL-2-R expression on both splenic B cells and PNA<sup>+</sup> thymocytes (12, 13).

In this study, we examined the effect of mIL-5 on the in vitro induction of LAK activity as a basic study for modified adoptive immunotherapy.

**Materials and Methods**

**Mice.** Male C57BL/6 mice, 6–8 wk old, were supplied by the Shizuoka Animal Center (Shizouka, Japan).

**Tumors.** 203 glioma is a malignant glioma cell line originally induced by 20-methylchol-antherene from the C57BL/6 strain. The tumor cells were maintained in a monolayer form in tissue culture with Dulbecco's modified MEM (Nissui, Japan) supplemented with 10% heat-inactivated FCS (Gibco Laboratories, Grand Island, NY) and antibiotics (penicillin G, 200 U/ml; streptomycin sulfate, 50 µg/ml). Methylcholanthrene-induced mastocytoma (P815), originating from DBA/2 mice, and Moloney leukemia virus-induced T cell lymphoma (YAC-1), originating from A/Sn mice, were maintained in RPMI 1640 (Nissui, Japan) with 10% heat-inactivated FCS and antibiotics.

**Lymphokines.** Human rIL-2 (TGP-3;  $4.2 \times 10^4$  U/mg protein) was kindly supplied by Takeda Chemical Industry, Ltd., Osaka, Japan.

rmIL-5 was the culture supernatant of X63-Ag8-653 myeloma cells transfected with mIL-5 complementary DNA (14, 15). Culture supernatant (SN) (1,200 U/ml) was filtered and aliquots were stored at  $-70^{\circ}\text{C}$ . Activity was determined by polyclonal IgM plaque-forming cell (PFC) assay using TRF-responding BCL<sub>1</sub> cells (16). This culture SN contained no IL-1, IL-2, IL-3, IL-4, IL-6, IL-7, or CSF (data not shown). Control culture SN was collected from the original myeloma cell (X63-Ag8-653).

**Generation of LAK Cells.** Spleens were removed aseptically and crushed with the hub of a syringe in RPMI 1640. Spleen cells were incubated with mIL-5 at a concentration of  $2 \times 10^6$  cells/ml for 3 d in the presence of IL-2 in RPMI 1640 with 10% heat-inactivated FCS and antibiotics. The culture was performed in 5-ml wells of a 6-well plate (Corning Glass Works, Corning, NY). The cells were harvested, washed twice, and used as effectors.

**Cytotoxicity Assay.** A 4-h  $^{51}\text{Cr}$  release assay in round-bottomed microtiter plates (Corning Glass Works) was used to measure cytotoxicity. Target cells were labeled with 0.1 mCi  $^{51}\text{Cr}$  (Amersham Corp., Arlington Heights, IL) for 60 min at  $37^\circ\text{C}$ . Effectors were coincubated with the targets at various E/T ratios for 4 h at  $37^\circ\text{C}$  in humidified air with 5%  $\text{CO}_2$ . The radioactive SNs were harvested and counted with a gamma counter. The cytotoxic activity was calculated from the average of triplicate cultures as: percent of cytolysis =  $100 \times (\text{experimental release} - \text{spontaneous release}) / (\text{maximum release} - \text{spontaneous release})$ . Spontaneous release was determined by adding culture medium only to the target cells and maximum release was counted in detergent. All determinations in each experiment were performed in triplicate.

### Results

**The Effect of rIL-5 on the Induction of LAK Cells.** Freshly isolated spleen cells were incubated with IL-2, rIL-5, or a combination of IL-2 and rIL-5. As shown in Fig. 1, rIL-5 enhanced IL-2-induced LAK activity, although rIL-5 alone induced no cytotoxic activity. Control SN obtained from X63·Ag8.653 myeloma cells did not show such enhancement (data not shown). Freshly isolated spleen cells showed slight cytotoxicity against YAC-1, but little cytotoxicity against P815 and 203-glioma (data

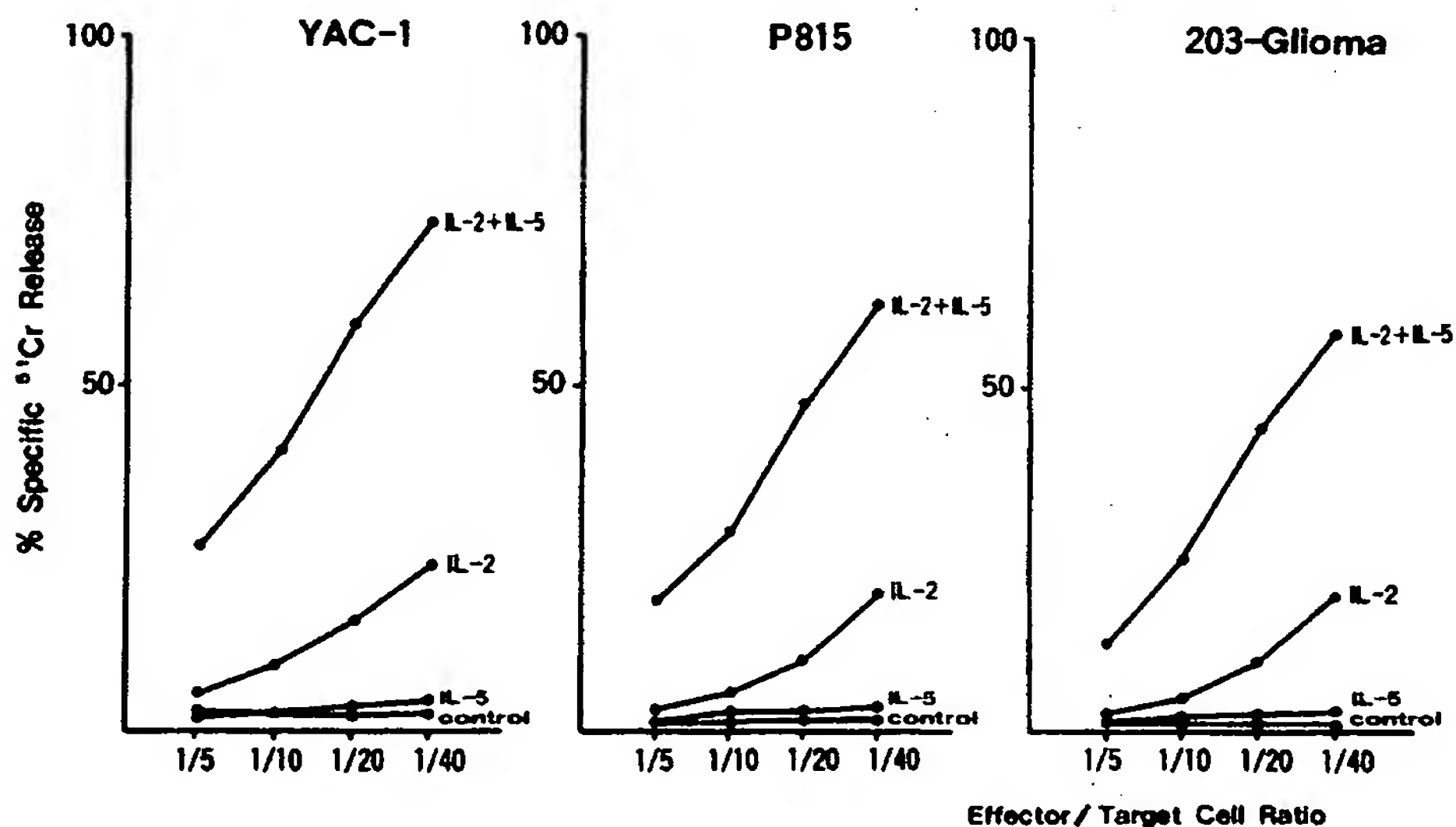


FIGURE 1. rIL-5 enhances rIL-2-mediated LAK activity. C57BL/6 mouse splenocytes were cultured at  $2 \times 10^6$  cells/ml in 5 ml of complete medium containing rIL-2 (1 U/ml) and/or rIL-5 (50 U/ml). After 3 d, effectors were tested for cytotoxicity against  $^{51}\text{Cr}$ -labeled tumor cells in a 4-h assay. Control values represent the killer activity of effectors cultured in complete medium without any lymphokines. Results are presented as the means of at least three experiments.

not shown). Spleen cells cultured without lymphokines for 3 d were ineffective in lysing YAC-1, P815, and 203-glioma.

*The Effect of rIL-5 on the Induction of LAK Activity Depends on the Concentration of IL-2.* IL-2 mediated LAK cell activity is also dependent on IL-5 concentration. To determine the optimal concentration of IL-2 or rIL-5 to obtain the maximum cytotoxic activity, double titration experiments were carried out. In Fig. 2, these killer activities show dose dependency at IL-2 concentrations from 0.01 to 10 U/ml and at rIL-5 concentrations from 0.01 to 100 U/ml.

*The Effects of Various Time Schedules of Administration of IL-2 and rIL-5 on the Induction of LAK Activity.* As shown in Fig. 3, rIL-5 alone did not show killer activity when IL-2 was not present from the onset of culture (Fig. 3, *j* and *k*). If IL-2 was present from the onset of culture, however, rIL-5 enhanced LAK activity, even when added 24–48 h after culture initiation (Fig. 3, *h* and *i*). Finally, the presence of IL-2 for the first 24 or 48 h and rIL-5 during the last 48 or 24 h, respectively, did not enhance LAK activity (Fig. 3, *d* and *e*). These results suggest that rIL-5 acts on the late stage of the LAK induction in the presence of IL-2.

### Discussion

The present study demonstrates that mIL-5 enhances the LAK activity in the presence of IL-2, though mIL-5 alone can not induce killer activity (Fig. 1). We also found that in vitro the cell numbers are greater from cultures containing IL-2 plus mIL-5 than those containing IL-2 alone (data not shown). As mIL-5 induces the expression of IL-2-R on both splenic B cells and PNA<sup>+</sup> thymocytes (12, 13), it may also induce the expression of IL-2-R on LAK cells. Although mIL-5 alone can

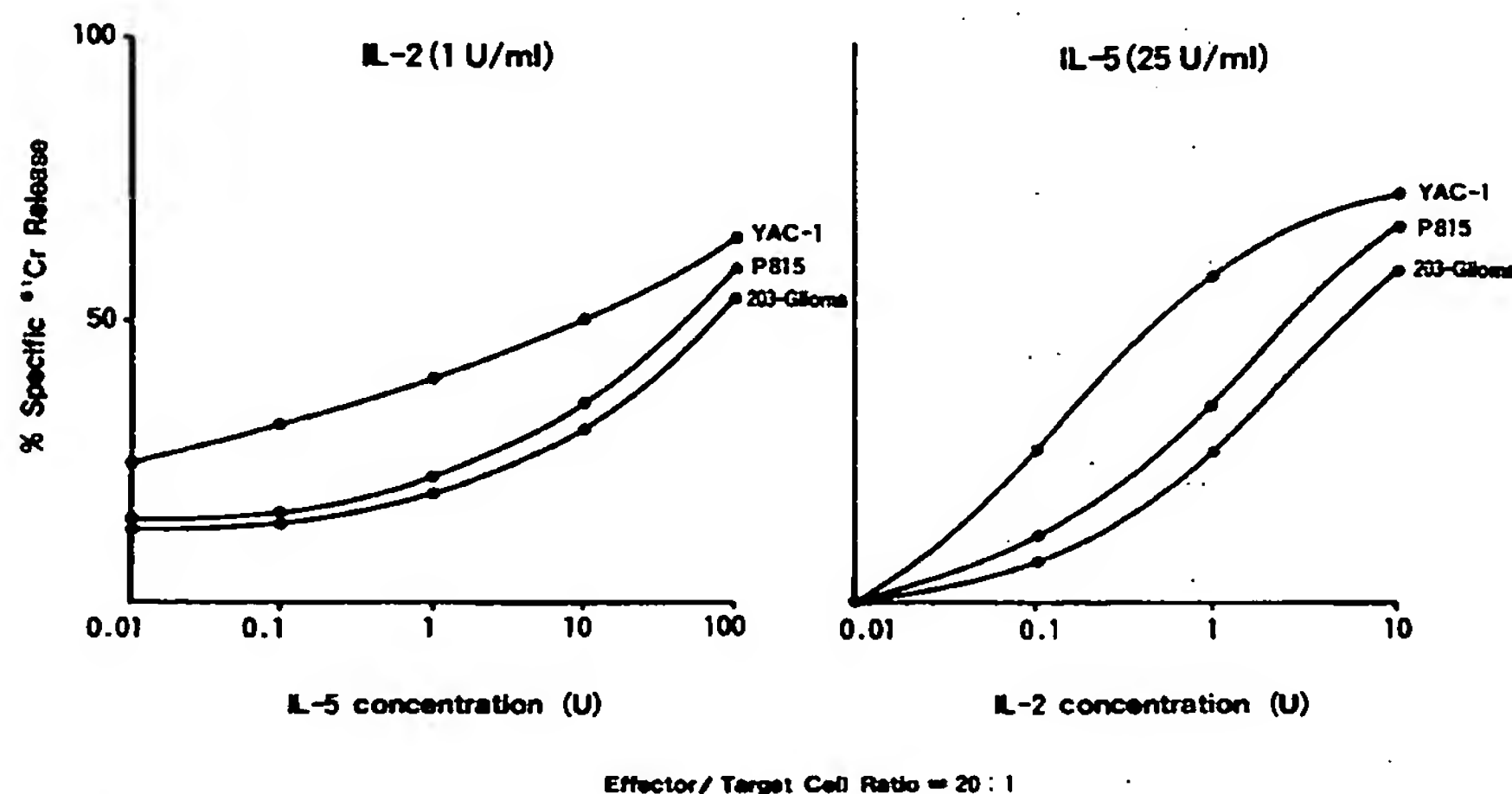


FIGURE 2. Generation of LAK activity. Dose-response cytotoxic activity by rIL-5 and rIL-2. C57BL/6 mouse splenocytes were cultured at various concentrations of rIL-5 and rIL-2 for 3 d. Cytotoxic activities against YAC-1, P815, and 203-Glioma were measured at an E/T ratio of 20:1. All determinations in each experiment were performed in triplicate.



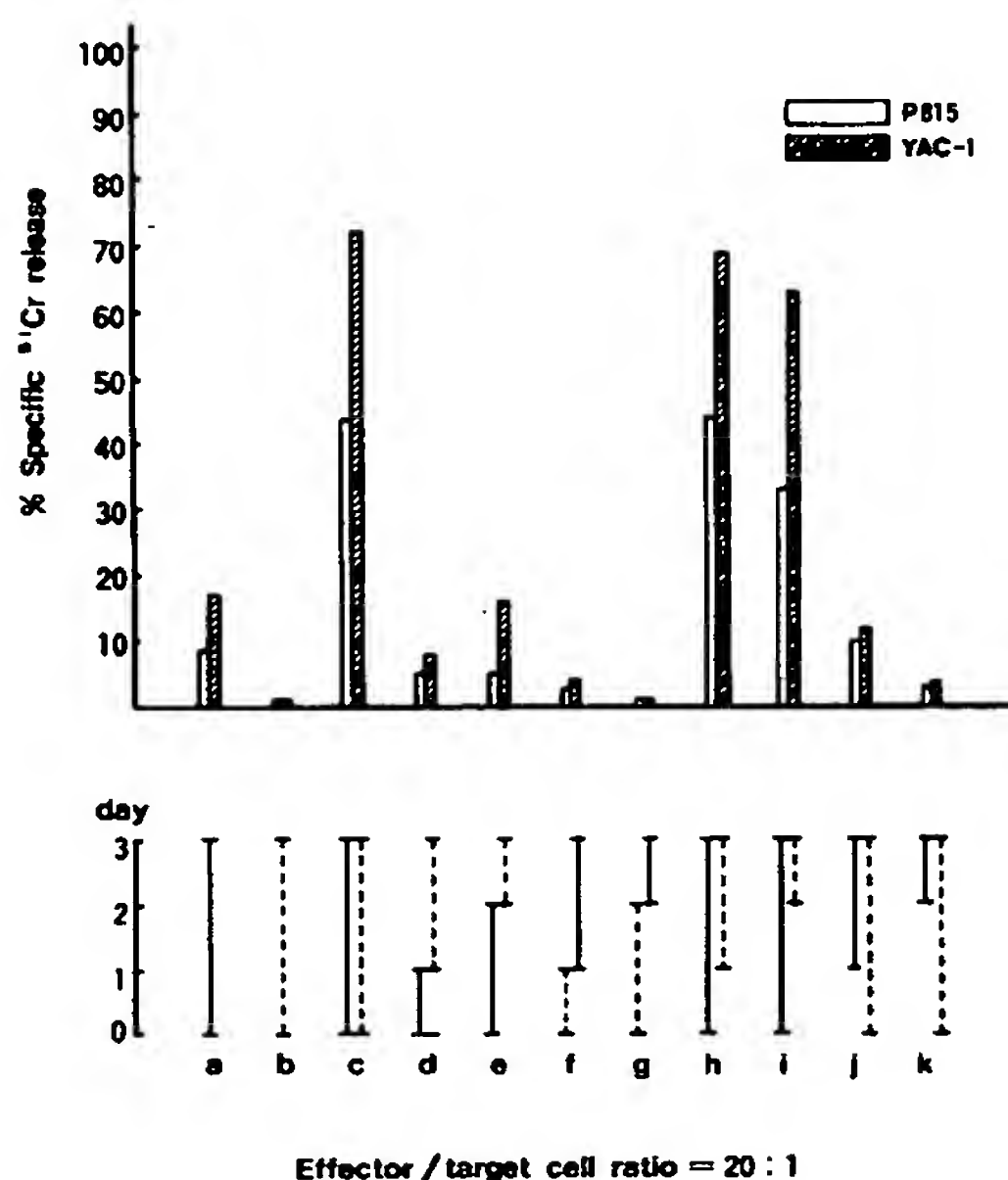


FIGURE 3. Effect of various time schedules of administration of rIL-5 and rIL-2 on LAK induction. C57BL/6 mouse splenocytes were cultured. rIL-5 (25 U/ml) or rIL-2 (1 U/ml) was present in the culture for limited periods as indicated. After 3 d, cytotoxic activities against YAC-1 and P815 were measured at an E/T ratio of 20:1. All determinations in each experiment were performed in triplicate. (—) IL-2; (-----) IL-5.

not induce killer activity, some mechanisms involving IL-5-R probably may augment IL-2 signal transduction pathways.

mIL-5 may act on the late stage of LAK induction in the presence of IL-2. As shown in Fig. 3, *c*, *h*, and *i*, if IL-2 is present from the onset of culture, mIL-5 enhances LAK activity, even when mIL-5 is added at the late stage of culture. If mIL-5 is present from the onset of culture, however, killer activity is not so enhanced, even when IL-2 is added at the late stage (Fig. 3, *j* and *k*). Hence, it might be possible that IL-2 activates precursor cells in the first stage, and mIL-5 augments the activated precursor cells. Moreover, this effect cannot take place in the absence of IL-2 (Fig. 3, *f* and *g*).

Mule et al. (17) found that mIL-4 alone induces killer activity as well as IL-2-mediated killer activity. mIL-5, on the other hand, enhances IL-2-mediated killer activity, but mIL-5 alone does not induce the killer activity. These results suggest that the mechanism by which LAK activity is augmented by mIL-5 is different from that of mIL-4.

Unlike IL-2, mIL-4 has little capacity to generate LAK activity from mouse PBL, whereas both lymphokines induce this activity from splenocytes (17). Whether or not mIL-5 and IL-2 similarly generate LAK activity from mouse PBL remains to be determined.

Hergen et al. (18) found that human IL-4 blocks IL-2-generated LAK activity. The reason for the discrepancy between human and mouse IL-4 is unclear. Apart from the possibility that the ability of IL-4 to block the induction of LAK activity is species restricted, it might be possible that mouse spleen cells contain already acti-

vated precursor cells. Hergen et al. (18) found IL-4 inhibits IL-2-induced LAK activity in spleen cells.

It would be interesting, for the purpose of adoptive immunotherapy, to see whether human IL-5 also enhances IL-2-induced LAK cell activity in PBL or not.

### Summary

IL-5 expresses various biologic effects on several types of lymphocytes, including B cells, eosinophils, and T cells. We demonstrated that the incubation of resting splenocytes from C57BL/6 mice in murine rIL-5 enhances IL-2-mediated lymphokine-activated killer (LAK) activity against various tumor cells. IL-5 alone, however, does not induce killer activity. IL-2-mediated LAK activity increases in proportion to the dose of IL-5. During the late phase of the culture period, IL-5 seems to have some effect on the induction of IL-2-mediated LAK activity. We expect that IL-5 will prove useful for adoptive immunotherapy.

*Received for publication 24 March 1989 and in revised form 4 May 1989.*

### References

1. Lots, M. T., E. A. Grimm, A. Mazumder, J. L. Strusser, and S. A. Rosenberg. 1981. Lysis of fresh and cultured autologous tumor by human lymphocytes cultured in T-cell growth factor. *Cancer Res.* 41:4420.
2. Grimm, E. A., A. Mazumder, H. Z. Zhang, and S. A. Rosenberg. 1982. Lymphokine-activated killer cell phenomenon. Lysis of natural killer-resistant fresh solid tumor cells by interleukin 2-activated autologous human peripheral blood lymphocytes. *J. Exp. Med.* 155:1823.
3. Rosenstein, M., I. Yron, Y. Kaufman, and S. A. Rosenberg. 1984. Lymphokine-activated killer cells: lysis of fresh syngeneic natural killer-resistant murine tumor cells by lymphocytes cultured in interleukin 2. *Cancer Res.* 44:1946.
4. Damle, N. K., L. V. Doyle, and E. C. Bradley. 1986. Interleukin 2-activated human killer cells are derived from phenotypically heterogeneous precursors. *J. Immunol.* 137:2814.
5. Phillips, J. H., and L. L. Lanier. 1986. Dissection of the lymphokine-activated killer phenomenon. Relative contribution of peripheral blood natural killer cells and T lymphocytes to cytotoxicity. *J. Exp. Med.* 164:814.
6. Itoh, K., A. B. Tilden, K. Kumagai, and C. M. Balch. 1985. Leu-11<sup>+</sup> lymphocytes with natural killer activity are precursors of recombinant interleukin 2-induced activated killer cells. *J. Immunol.* 134:802.
7. Takatsu, K., A. Tomizuka, and T. Hamaoka. 1980. Antigen-induced T cell-replacing factor (TRF). I. Functional characterization of helper T lymphocytes and genetic analysis of TRF production. *J. Immunol.* 124:2414.
8. Tominaga, A., K. Takatsu, and T. Hamaoka. 1980. Antigen-induced T cell-replacing factor (TRF). I. Establishment of T cell hybrid clone continuously producing TRF and functional analysis of released TRF. *J. Immunol.* 126:2646.
9. Swan, S. L., and R. Dutton. 1982. Production of a B cell growth-promoting activity, (DL) BCGF, from a cloned T cell line and its assay on the BCL<sub>1</sub> B cell tumor. *J. Exp. Med.* 156:1821.
10. Yamaguchi, Y., T. Suda, M. Eguchi, Y. Mimura, N. Harada, A. Tominaga, and K. Takatsu. 1988. Purified interleukin (IL-5) support the terminal differentiation proliferation of murine eosinophilic precursors. *J. Exp. Med.* 167:43.

11. Takatsu, K., Y. Kikuchi, T. Takahashi, T. Honjo, M. Matsumoto, N. Harada, N. Yamaguchi, and A. Tominaga. 1987. Interleukin 5, a T-cell-derived B-cell differentiation factor also induces cytotoxic T lymphocytes. *Proc. Natl. Acad. Sci. USA.* 84:4234.
12. Loughan, M. S., K. Takatsu, N. Harada, and G. J. V. Nossal. 1987. T-cell-replacing factor (interleukin 5) induces expression of interleukin 2 receptors on murine splenic B cells. *Proc. Natl. Acad. Sci. USA.* 84:5399.
13. Yuji, K., R. Kato, Y. Sano, H. Takahashi, T. Kanatani, and K. Takatsu. 1986. Generation of cytotoxic T lymphocytes from thymocyte precursors to trinitrophenyl modified self antigens. *J. Immunol.* 136:1161.
14. Kinashi, T., N. Harada, E. Severinson, T. Tanabe, P. Sideras, C. Azuma, A. Tominaga, S. Bergstedt-Lindqvist, M. Takashi, M. Konishi, F. Matsuda, Y. Yokota, K. Takatsu, and T. Honjo. 1986. Cloning of complementary DNA encoding T-cell replacing factor and identity with B-cell growth factor I. *Nature (Lond.).* 324:70.
15. Karasuyama, H., and F. Melchers. 1988. Establishment of mouse cell lines which constitutively secrete large quantities of interleukin 2, 3, 4, or 5 using modified cDNA expression vectors. *Eur. J. Immunol.* 18:97.
16. Takatsu, K., N. Harada, Y. Takahara, G. Yamada, K. Dobashi, and T. Hamaoka. 1985. Purification and physicochemical characterization of murine T cell replacing factor (TRF). *J. Immunol.* 134:382.
17. Mule, J. J., C. A. Smith, and S. A. Rosenberg. 1987. Interleukin 4 (B cell stimulatory factor 1) can mediate the induction of lymphokine-activated killer cell activity directed against fresh tumor cells. *J. Exp. Med.* 166:792.
18. Hergen, S., H. Yssel, X. Paliard, R. Kastelein, C. Figdor, and J. Vries. 1988. IL-4 inhibits IL-2-mediated induction of human lymphokine activated killer cells, but not the generation of antigen specific cytotoxic T lymphocytes in mixed leucocytes. *J. Immunol.* 141:29.

# blood

1994 84: 2261-2268

## **Interleukin-12 is chemotactic for natural killer cells and stimulates their interaction with vascular endothelium**

P Allavena, C Paganin, D Zhou, G Bianchi, S Sozzani and A Mantovani

---

Information about reproducing this article in parts or in its entirety may be found online at:  
[http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#repub\\_requests](http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#repub_requests)

Information about ordering reprints may be found online at:  
<http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#reprints>

Information about subscriptions and ASH membership may be found online at:  
<http://bloodjournal.hematologylibrary.org/site/subscriptions/index.xhtml>





## Interleukin-12 Is Chemotactic for Natural Killer Cells and Stimulates Their Interaction With Vascular Endothelium

By Paola Allavena, Carla Paganin, Dan Zhou, Giancarlo Bianchi, Silvano Sozzani, and Alberto Mantovani

We investigated the chemotactic activity of interleukin (IL)-12 on human natural killer (NK) cells and other leukocyte subsets. It was found that IL-12 induced directional migration of highly enriched preparations of NK cells (>80% CD16<sup>+</sup> and CD56<sup>+</sup>) and CD3-activated T cells (both of CD4 and CD8 subset), but not resting T cells and monocytes. On the contrary, purified polymorphonuclear cells (PMN) showed significant and reproducible chemotactic response to IL-12. The effects of IL-12 on leukocyte migration were observed in a narrow concentration range with a peak at approximately 7.5 ng/mL, and were abrogated by monoclonal antibody (MoAb) anti-IL-12 or after cytokine boiling. We also investi-

gated the interaction of NK cells with vascular endothelium in vitro. Overnight treatment of NK cells with IL-12 augmented their binding to cultured endothelial cells (EC) obtained from umbilical veins. IL-12-increased binding was better observed when resting rather than IL-1-activated EC were used as substratum of adhesion. IL-12-augmented binding of NK cells to resting or IL-1-activated EC involved the LFA-1/ICAM-1, and VLA-4/VCAM-1 pathways. Thus, by inducing migration and interaction with EC, IL-12 regulates crucial determinants of NK-cell recruitment in tissues.  
© 1994 by The American Society of Hematology.

**N**ATURAL KILLER (NK) cells comprise 5% to 10% of blood mononuclear cells and are characterized by the typical morphology of large granular lymphocytes (LGLs) and the CD3<sup>+</sup>CD16<sup>+</sup>CD56<sup>+</sup> phenotype. NK cells are cytotoxic without prior sensitization for a wide range of target cells, including some microorganisms and virus-infected and transformed cells. In addition to cytotoxicity, NK cells secrete cytokines and have a central role in the regulation of the immune response and hematopoiesis.<sup>1-3</sup> Most NK cells are found mainly in blood and spleen, but they are present also in the lung and intestinal mucosa, and accumulate in liver parenchyma on cytokine administration in mice.<sup>2</sup> In certain pathologic conditions, NK cells accumulate at extrahematic sites following viral and bacterial infections,<sup>4,5</sup> and are found in the allograft infiltrate during the early phase of rejection.<sup>6</sup>

The mechanisms underlying NK-cell recruitment from blood vessels to tissues have not been fully elucidated. We and others have reported that NK cells are rapid mobile cells during in vitro chemotaxis assays,<sup>7-10</sup> express adhesion molecules, which can recognize ligands on resting and inflamed endothelium<sup>11,12</sup> or extracellular matrix proteins,<sup>13,14</sup> and are able to actively transmigrate across endothelial monolayers.<sup>15</sup>

These functions are modulated by cytokines. Interleukin-2 (IL-2), interferon (IFN)- $\gamma$  and tumor necrosis factor (TNF)- $\alpha$  stimulate their in vitro locomotor ability, and recently it was reported that IL-8 is chemokinetic for IL-2-activated adherent NK cells.<sup>16-18</sup> The adhesive interaction of NK cells to in vitro-cultured endothelial cells (EC) is upregulated by IL-2,<sup>12,19</sup> while IL-4 partly inhibited their binding.<sup>20</sup>

In the present work, we studied IL-12, a cytokine produced by B lymphocytes and monocytes in response to lipopolysaccharide (LPS) and bacterial products, recently cloned and characterized.<sup>21-23</sup> IL-12 has important functions in the regulation and amplification of the immune response and inflammation.<sup>24</sup> IL-12 regulates the cytotoxic activity, proliferation, and cytokine production of NK and T cells,<sup>21,25,26</sup> but no information is available on modulation of cell adherence or chemotaxis.

Our results show that IL-12 is chemotactic for NK cells and increases their in vitro binding to and killing of EC.

Further, IL-12 induced directional migration of CD3-activated T lymphocytes and polymorphonuclear cells (PMN), but not resting T cells or monocytes.

### MATERIALS AND METHODS

**Culture media and reagents.** The following reagents were used for separation of effector cells, cell culture, and experimental assays: pyrogen-free saline for clinical use and distilled water (Bieffe, Sondrio, Italy); medium RPMI 1640 (10 $\times$  concentrated; Biochrom KG, Berlin, Germany); medium 199 (GIBCO, Paisley, UK); glutamine (GIBCO); penicillin and streptomycin (GIBCO); and aseptically collected fetal calf serum (FCS; Hyclone, Logan, UT). The routinely used tissue culture medium was RPMI 1640 with 2 mmol/L glutamine, 50 mg/mL gentamicin, and 10% FCS, which is hereafter referred to as complete medium.

All reagents were checked for endotoxin contamination by the Limulus amoebocyte lysate assay (Microbiological Associates, Walkersville, MD) with a sensitivity of 0.02 to 0.05 ng/mL *Escherichia coli* Westphal LPSs. Sera were tested after 1:3 dilution and heating at 100°C. All reagents were negative for endotoxin contamination. LPS (*E coli* 0.55:B5) was purchased from Difco (Detroit, MI).

Recombinant C5a (a kind gift from Dr V. Showell, Pfizer Control Research, Groton, CT) was used as reference chemoattractant for the different leukocyte populations. IL-12 was a kind gift of Genetics Institute, Cambridge, MA. IL-2 was a kind gift of Roussel Uclaf, Paris, France.

---

*From the Department of Immunology, "Mario Negri" Institute, Milano, Italy.*

*Submitted January 10, 1994; accepted June 6, 1994.*

*Supported by a contract from the National Program of Pharmacological Research (Rif.1146/193/07/8602), by the Italian Consortium for Immunomodulation Technologies (C.I.T.I.) for the Italian Ministry of University and Technological Research; by the Program Italy-USA on Cancer Therapy, by Consiglio Nazionale delle Ricerche (finalized project ACRO), and by the Italian Association for Cancer Research (AIRC).*

*Address reprint requests to Paola Allavena, MD, Laboratory of Immunology, "Mario Negri" Institute, Via Eritrea, 62-20157 Milano, Italy.*

*The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. section 1734 solely to indicate this fact.*

© 1994 by The American Society of Hematology.  
0006-4971/94/8407-0007\$3.00/0

**Preparation of leukocyte populations.** Leukocytes were obtained from buffy coats of blood donors from normal volunteers through the courtesy of Centro Transfusionale, Ospedale Sacco, Milan, Italy, by centrifugation on Ficoll-Hypaque gradient. PMN were isolated as described.<sup>27</sup> Briefly, cells from the pellet of the Ficoll-Hypaque gradient (Biochrom KG, Berlin, Germany) were resuspended in isoosmotic (285 mOsm) complete medium, layered onto 62% isoosmotic Percoll gradient (diluted with complete medium), and spun at 450g for 20 minutes. Cells were collected at the medium/Percoll interface. The purity of PMN preparations was more than 95%, as assessed morphologically on Giemsa-stained cytopreps. Monocytes were separated on 46% Percoll gradients as described.<sup>28</sup> Preparations were usually more than 85% pure as assessed by positivity with monoclonal antibody (MoAb) anti-CD14.

Highly enriched NK-cell populations were prepared by discontinuous (47%, 49%, and 52%) Percoll gradients as previously described.<sup>7</sup> Low-density cells were further depleted of contaminating T lymphocytes by panning with MoAb anti-CD6 (50 ng/mL).<sup>15</sup> Cells from NK-enriched fractions were added to the dishes and incubated at 4°C for 1 hour. The nonadherent cells were then gently poured off. The resulting NK populations were less than 3% CD3<sup>+</sup> and more than 80% CD16<sup>+</sup> and CD56<sup>+</sup>.

T lymphocytes were obtained from the high-density fractions of discontinuous Percoll gradients and further purified by positive panning on CD6-coated petri dishes. In some experiments, T cells were separated in CD4 and CD8 subsets by panning. For activation with anti-CD3, T cells were treated with a saturating amount of anti-CD3 MoAb and incubated on a plastic Petri dish previously coated with affinity-purified goat-antimouse IgG (10 µg/mL; Sigma Chemical Co, St Louis, MO) for 1 hour at 4°C. Nonadherent cells were gently poured off and cell incubation was continued at 37°C for an additional 4 hours.

**Migration assay.** Cell migration was evaluated using a micro-chamber technique<sup>29</sup> as previously described for monocytes and PMN,<sup>27</sup> and for NK cells and T cells.<sup>7</sup> Twenty-five ± 1 µL chemoattractant diluted in RPMI with 1% FCS was seeded in the lower compartment of the chemotaxis chamber and 50 µL cell suspension (1.5 × 10<sup>6</sup>/mL monocytes in peripheral blood mononuclear cells (PBMC) or PMN or 2 × 10<sup>6</sup>/mL LGL or T lymphocytes) were seeded in the upper compartment. The two compartments were separated by 8-µm pore size nitrocellulose or polyvinylpyrrolidone-free polycarbonate membranes (Nucleopore Corp, Pleasanton, CA). Chambers were incubated at 37°C in air with 5% CO<sub>2</sub> for 1.5 hours (monocytes), 2 hours (PMN), 4 hours (LGL), or 3 hours (T cells). At the end of the incubation, nitrocellulose filters were removed, fixed, stained, and dehydrated using standard histologic methods. Five fields were examined for each sample. Migration was expressed as the distance (in microns) migrated by the two leading cells. Expected values of migration, assuming no chemotactic response, were calculated as described by Zigmond and Hirsch.<sup>30</sup> Polycarbonate filters were fixed and stained with Diff-Quik (Baxter, Dürdinen, Germany) and five high-power oil fields were counted. Migration was expressed as the number of cells migrated across the filter.

**Preparation of EC.** Human EC were obtained from umbilical vein and cultured as previously described.<sup>31</sup> Routinely, we used confluent cells (10<sup>5</sup>/2 cm<sup>2</sup> culture well) between the first and fourth passage maintained in 199 medium with 20% bovine serum (Hyclone), supplemented with EC growth supplement (50 µg/mL; Collaborative Research Inc, Lexington, MA) and heparin (100 µg/mL; Sigma). The purity of EC cultures was checked by expression of von Willebrand factor and found to be more than 99% positive.

**Antibodies.** The following MoAbs were used: anti-CD18 (clone TS1/18), anti-CD11a (clone TS1/22), and anti-CD6 (clone 3P12) purchased from ATCC (Rockville, MD); anti-CD11a (clone NKI-

L16) kindly obtained from Dr C Figdor (Nijmegen, The Netherlands); anti-CD11b (clone 44a) kindly obtained from Dr R. Todd (Ann Arbor, MI); anti-CD11c (clone L29) (Becton Dickinson, Mountain View, CA) kind gift of Dr L. Lanier (DNAX, Palo Alto, CA); anti-VLA4 (clone HP2/1) kindly obtained from Dr F. Sanchez-Madrid, (University of Madrid, Madrid, Spain); anti-CD3 (clone OKT3) kind gift of Dr S. Ferrini (IST, Genova, Italy); anti-CD16 (clone B73.1) and anti-IL-12, kind gifts of Dr G. Trinchieri (Wistar Institute, Philadelphia, PA); anti-CD56 (clone N901) kind gift of Dr J. Griffin (Dana Farber, Boston, MA); and anti-CD14 (clone UCHM1) kind gift of Dr P. Beverly (University College and Middle Sex School of Medicine, London, UK). In the adhesion assays, MoAbs were used at a previously determined optimal concentration,<sup>12</sup> usually 1/100 for ascites or diluted 1/5 to 10 for hybridoma supernatants. NK cells were pretreated for 15 minutes at room temperature with MoAbs and then plated (without washing out the MoAb) in the adhesion assay. In each experiment, isotype-matched antibodies were used as negative control.

**Adhesion assay.** Adhesion of NK cells to EC was studied as described previously.<sup>12</sup> EC were grown to confluence in flat-bottomed 96-well trays and, in some experiments, were activated for 20 hours with 10 ng/mL of IL-1β (Dompè, L'Aquila Italy).

NK-cell preparations were resuspended at 2 × 10<sup>6</sup>/mL in complete medium and incubated overnight with the stimulating cytokine. Control NK cells were incubated in medium alone. At the end of the culture period, cells were washed with saline and resuspended in complete medium at 10<sup>7</sup>/mL for labeling with 100 µCi <sup>51</sup>Cr (sodium chromate; Amersham, Bucks, UK) at 37°C for 1 hour. Cells were then washed with medium and resuspended at 10<sup>6</sup>/mL. 0.1 mL was dispensed to each 96-well tray and incubated for 30 minutes at 37°C. At the end of the incubation, the wells were carefully washed three times with phosphate-buffered saline (PBS) plus 1% fetal bovine serum; adherent cells were solubilized with 0.2 mL of 0.025 mol/L NaOH and 0.1% sodium dodecyl sulfate (SDS), and radioactivity was counted in a gamma counter. Results are presented as the percentage of adherent cells ± SD, with 3 to 6 replicates per group. Statistical analysis was performed by the Mann-Whitney test.

**Cytotoxicity assay.** The cytotoxic assay was performed as described.<sup>21</sup> EC were labeled with 50 µL <sup>51</sup>Cr (Amersham) for 1 hour, and 5 × 10<sup>3</sup> EC were cocultured with a different concentration of NK cells previously treated with IL-2 or IL-12 for 24 hours. After a 4-hour incubation, supernatants were harvested and counted in a gamma counter for determination of isotope release. Results are presented as percent specific lysis (mean ± SD) of three replicates per group.

## RESULTS

**Chemotactic activity of IL-12.** In the first series of experiments, we tested the chemotactic activity of IL-12 on purified NK cells and for comparison on other leukocyte populations.

Figure 1A shows a representative dose/response experiment. A chemotactic effect was observed in a narrow dose range, with a peak at 7.5 ng/mL of IL-12; the curve was typically bell-shaped, as for many chemoattractants. At the optimal concentration of 7.5 ng/mL, IL-12 was as active as an optimal concentration (50 ng/mL) of recombinant C5a. This result was confirmed in eight experiments with different blood donors. When an MoAb anti-IL-12 or boiled IL-12 were added in the assay, no migration occurred, confirming the specificity of the response (Fig 1B). This enhanced locomotor activity of NK cells was chemotactic rather than



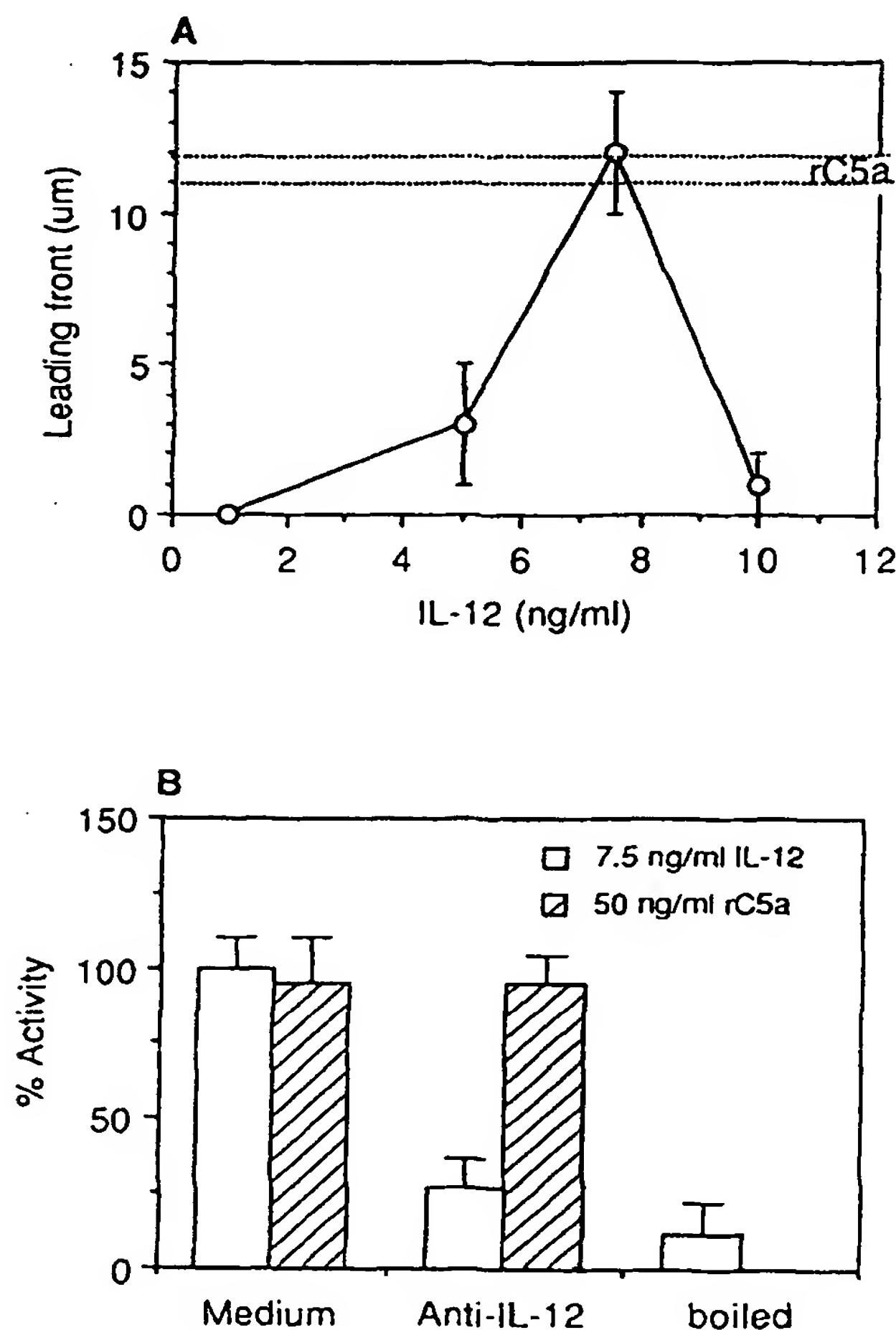


Fig 1. Chemotactic effect of IL-12 on highly enriched human NK-cell preparations. (A) Dose/response of IL-12. Recombinant C5a (50 ng/mL; ----) was used as reference chemoattractant. Results are expressed as migrated distance ( $\mu\text{m}$ ) of the two leading cells. (B) The chemotactic activity in response to IL-12 or C5a was assessed in the presence of anti-IL-12 MoAb or IL-12 boiled for 10 minutes. The migration in response to IL-12 alone or IL-12 in the presence of an irrelevant MoAb was considered as 100%.

chemokinetic, as demonstrated in checkerboard experiments, although a chemotactic component was also present (Table 1). Purified populations of T lymphocytes ( $>98\%$  CD3 $^{+}$ ) responded to IL-12 only when appropriately stimulated with anti-CD3 MoAb. Resting T cells (Fig 2A) or T cells incubated with an irrelevant MoAb (data not shown) did not respond. IL-12 was as potent as recombinant C5a in inducing T-cell migration. We also investigated whether the CD4 and CD8 subset differed in their chemotactic response to IL-12. Purified preparations of CD4 $^{+}$  and CD8 $^{+}$  T cells, preactivated with MoAb anti-CD3, showed similar chemotactic activity (Fig 2B).

While monocytes did not show increased locomotor activity in response to IL-12 (data not shown), surprisingly, purified preparations of PMN showed a significant and reproduc-

ible chemotactic response to IL-12, with a peak effect at 5 ng/mL. Figure 3A shows a typical experiment representative of seven performed. IL-12 activity on PMN was blocked by anti-IL-12 MoAb, and the heat-treated cytokine was ineffective (Fig 3B). A checkerboard analysis was also performed for PMN, and it was found that IL-12 is chemotactic, although a small chemokinetic effect was also evident (data not shown).

*IL-12 stimulates adhesion and cytotoxicity of NK cells to vascular endothelium.* The effect of IL-12 was also studied on the ability of NK cells to adhere in vitro to monolayers of cultured human EC. IL-12 increased NK-cell binding to resting, as well as to IL-1-activated EC, as shown in the representative experiment in Fig 4, although IL-2, used in parallel as an NK-cell-activating cytokine, was a stronger stimulus compared with IL-12. This IL-12-increased binding was better observed when resting rather than activated EC were used as substratum for adhesion. In a large series of experiments, the estimated increase in NK-cell binding induced by IL-12 was  $68\% \pm 10\%$  and  $33\% \pm 3\%$  on resting and IL-1-activated EC, respectively (mean  $\pm$  SE of 15 experiments,  $P < .05$ ).

As shown in Fig 5, a small increase in NK-cell binding to EC was already observed with 5 ng/mL IL-12, and the peak effect was at 10 ng/mL. These concentrations were also the most effective in parallel experiments of tumor cytotoxicity against the K562 cell line (Fig 5). The effect of IL-12 required an incubation time of more than 6 hours, and was optimal after an overnight incubation (not shown). No increase in NK-cell binding was observed when EC were stimulated for 24 hours with IL-12 (data not shown).

As IL-12-induced stimulation of binding to EC required a relatively long exposure, we questioned whether a putative factor secreted by NK cells was, in fact, responsible for the increased binding. IL-12 induces the production of several cytokines by NK cells, eg, IFN $\gamma$  and TNF $\alpha$ .<sup>25</sup> These cytokines have stimulating effects on NK-cell function, but do not increase their binding ability to EC.<sup>20</sup> Moreover, MoAbs against TNF $\alpha$  and IFN $\gamma$  did not decrease NK-cell binding (data not shown). Also, the supernatant of NK cells cultured overnight with IL-12 did not stimulate NK-cell binding (not shown). These experiments suggested that IL-12 acts directly on the adhesive capacity of NK cells.

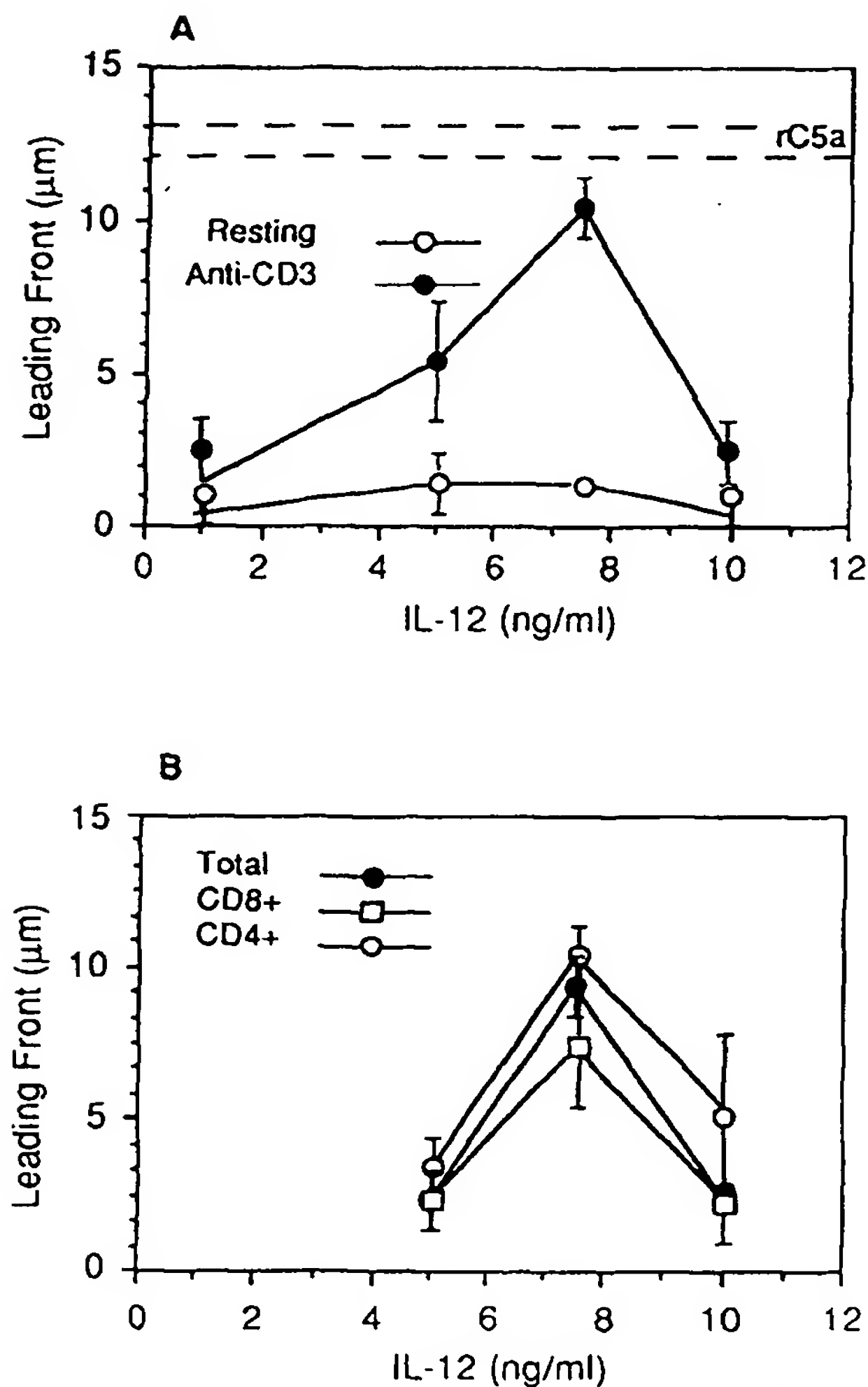
IL-2-activated NK cells are able to kill EC.<sup>20,32,33</sup> There-

Table 1. Checkerboard Analysis of IL-2 Stimulation of LGL Migration Across Nitrocellulose Filter

Below (ng/mL)	Medium	Above (ng/mL)		
		1	5	7.5
Medium	15 $\pm$ 0	18 $\pm$ 1 (16.9)	18 $\pm$ 1 (16.9)	19 $\pm$ 1* (16.9)
1	14 $\pm$ 0 (15.1)	17 $\pm$ 2	19 $\pm$ 1* (17.0)	19 $\pm$ 1* (17.0)
5	17 $\pm$ 1 (15.1)	17 $\pm$ 1 (17.0)	17 $\pm$ 2	19 $\pm$ 1* (17.0)
7.5	24 $\pm$ 1* (10.8)	21 $\pm$ 1* (17.0)	18 $\pm$ 2 (17.0)	17 $\pm$ 1

Different concentrations of IL-12 (ng/mL) were placed in the upper and/or lower compartments of the chemotaxis chamber. Migration is expressed as leading front distance ( $\mu\text{m} \pm$  SD). Numbers in parentheses are the expected values of migration calculated according to the method of Zigmond and Hirsch,<sup>20</sup> assuming no chemotactic response.

\*  $P < .05$  v migration to control medium (above and below the filter).



**Fig 2.** Chemotactic effect of IL-12 on purified T lymphocytes. (A) Dose/response of IL-12 with unstimulated and CD3-stimulated T cells. Results are expressed as migrated distance (μm) of the two leading cells. (B) T cells were separated in CD4<sup>+</sup> and CD8<sup>+</sup> cells by panning and stimulated with anti-CD3 before the assay. Chemotactic response to rC5a was  $12 \pm 2$  (μm) for unseparated (total) T cells,  $12 \pm 2$  for CD4<sup>+</sup>, and  $9 \pm 2$  for CD8<sup>+</sup>.

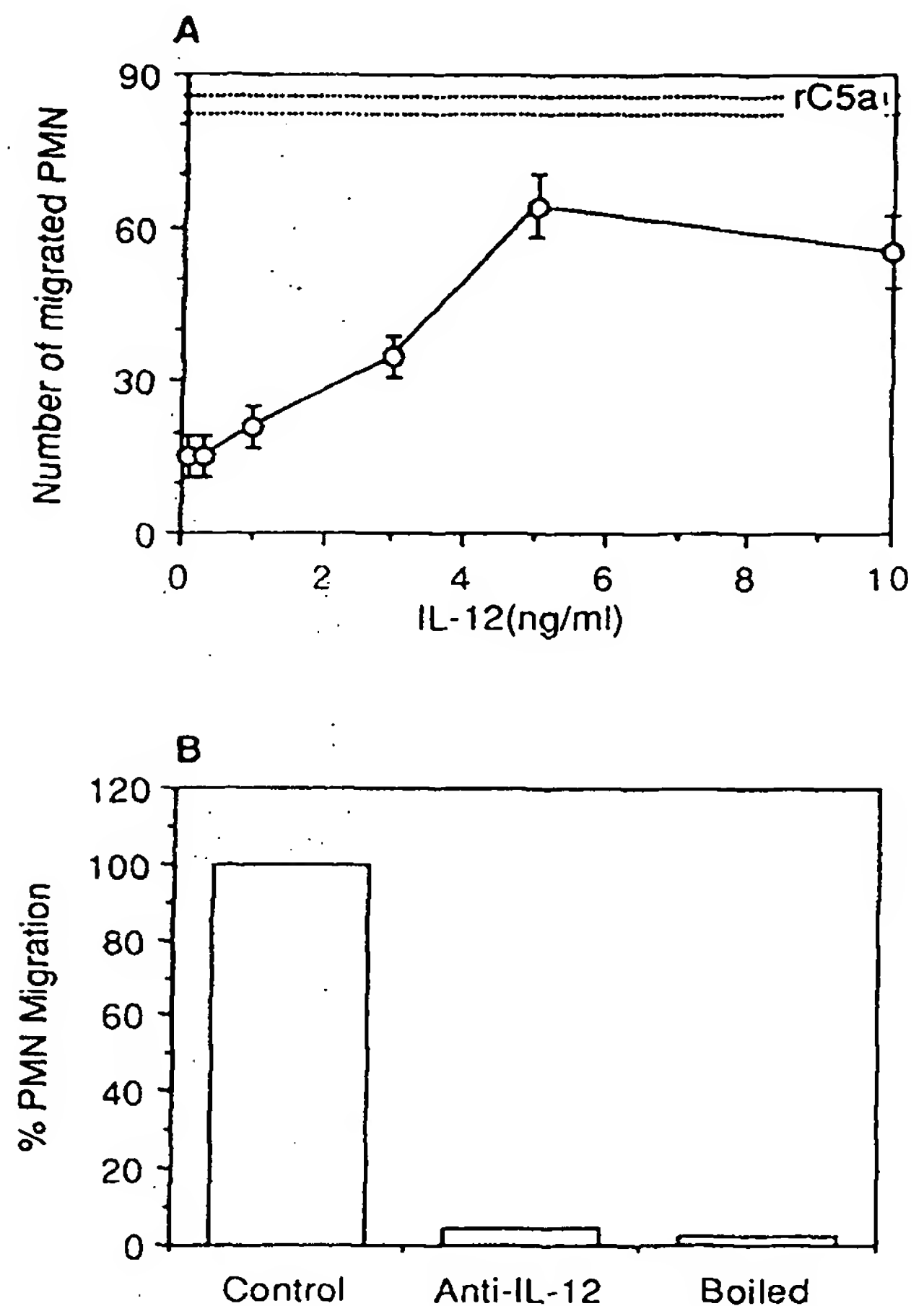
fore, we investigated whether the increased binding of IL-12-activated NK cells to EC also resulted in increased cytotoxicity. As shown in Fig 6, unstimulated peripheral blood lymphocytes (PBL) did not lyse EC; after overnight exposure to IL-12, a significant increase of lysis was detected, although lower than that obtained after IL-2 stimulation.

**Characterization of the adhesive molecules.** In previous reports, we characterized the adhesive molecules involved in NK-cell binding to EC.<sup>12</sup> Most binding of resting and IL-12-activated NK cells on resting EC was inhibited by anti-CD11a/CD18 MoAb, whereas MoAb anti-CD11b and CD11c had no effect (Fig 7). When EC were preactivated with IL-1, a clear role for the VLA-4/VCAM-1 pathway was

evident, as demonstrated by the fact that anti-VLA-4 MoAb used in concert with anti-CD18 gave a greater inhibition compared with anti-CD18 alone (Fig 7).

The observation that the greatest effect of IL-12 was seen on resting EC, where most binding is mediated by CD11a/CD18, caused us to check if this augmented functional activity was associated with an augmented expression of  $\beta_2$  integrins.

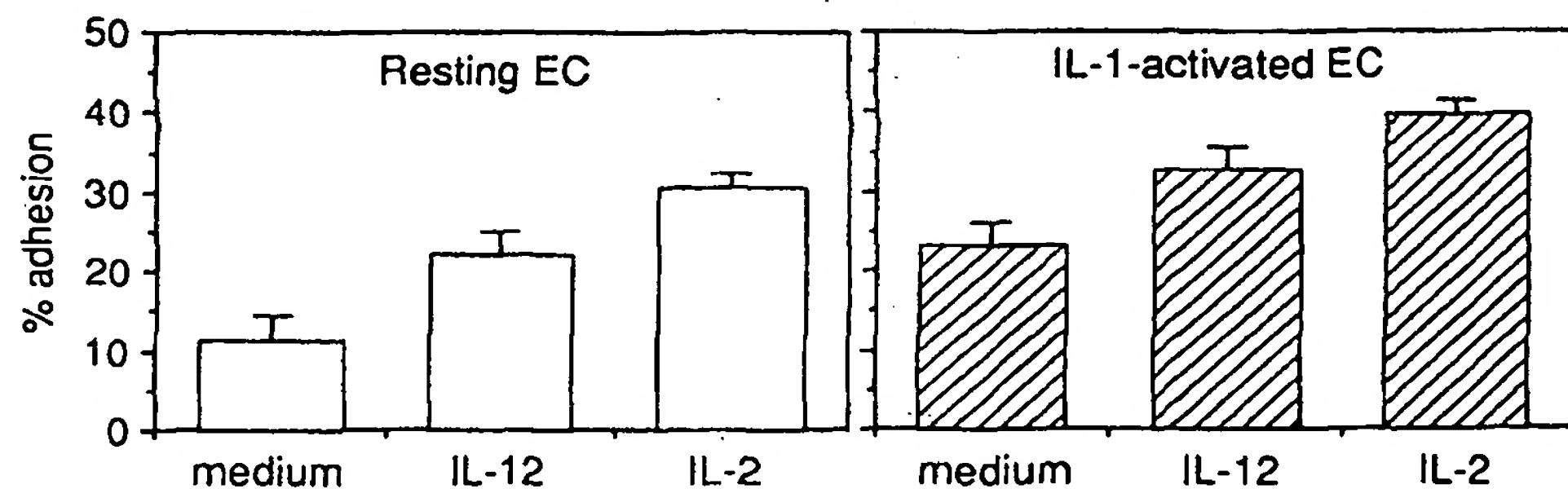
In several experiments, we were unable to demonstrate a consistent increased expression of  $\beta_2$  integrins after incubation with IL-12, using MoAb anti-CD11a (clone TS1/22) and anti-CD18 (clone TS1/18). When we used the MoAb NKI-L16, which recognizes an activation epitope of CD11a,<sup>34</sup> we did see an increased expression of NKI-L16 epitope after IL-12 stimulation (Fig 8). As MoAb anti-NKI-L16 stained already more than 80% of the cells, increased



**Fig 3.** Chemotactic effect of IL-12 on purified preparations of human PMN. (A) Dose/response of IL-12. Results are expressed as the number of PMN migrated through the filter. (B) PMN migration induced by IL-12 (5 ng/mL) in the presence of anti-IL-12 MoAb (ascites diluted 1:100) or by IL-12 boiled for 10 minutes. PMN migration in response to IL-12 alone or IL-12 plus an irrelevant MoAb was considered 100%.



**Fig 4.** Adhesion of highly enriched preparations of human NK cells to in vitro cultured human EC. NK cells were treated overnight with IL-12 (10 ng/mL), IL-2 (100 U/mL) or medium. Percent of adhesion by  $^{51}\text{Cr}$ -labeled NK cells was evaluated after 30 minutes incubation at 37°C. Values are mean  $\pm$  SD (four replicates) from one representative experiment.



expression is detected by mean channel fluorescence (eg, from 587 to 639). This result was confirmed in two additional experiments in which IL-12-treated cells (24 hours) showed an increase in fluorescence intensity of 32 and 40 channels, respectively.

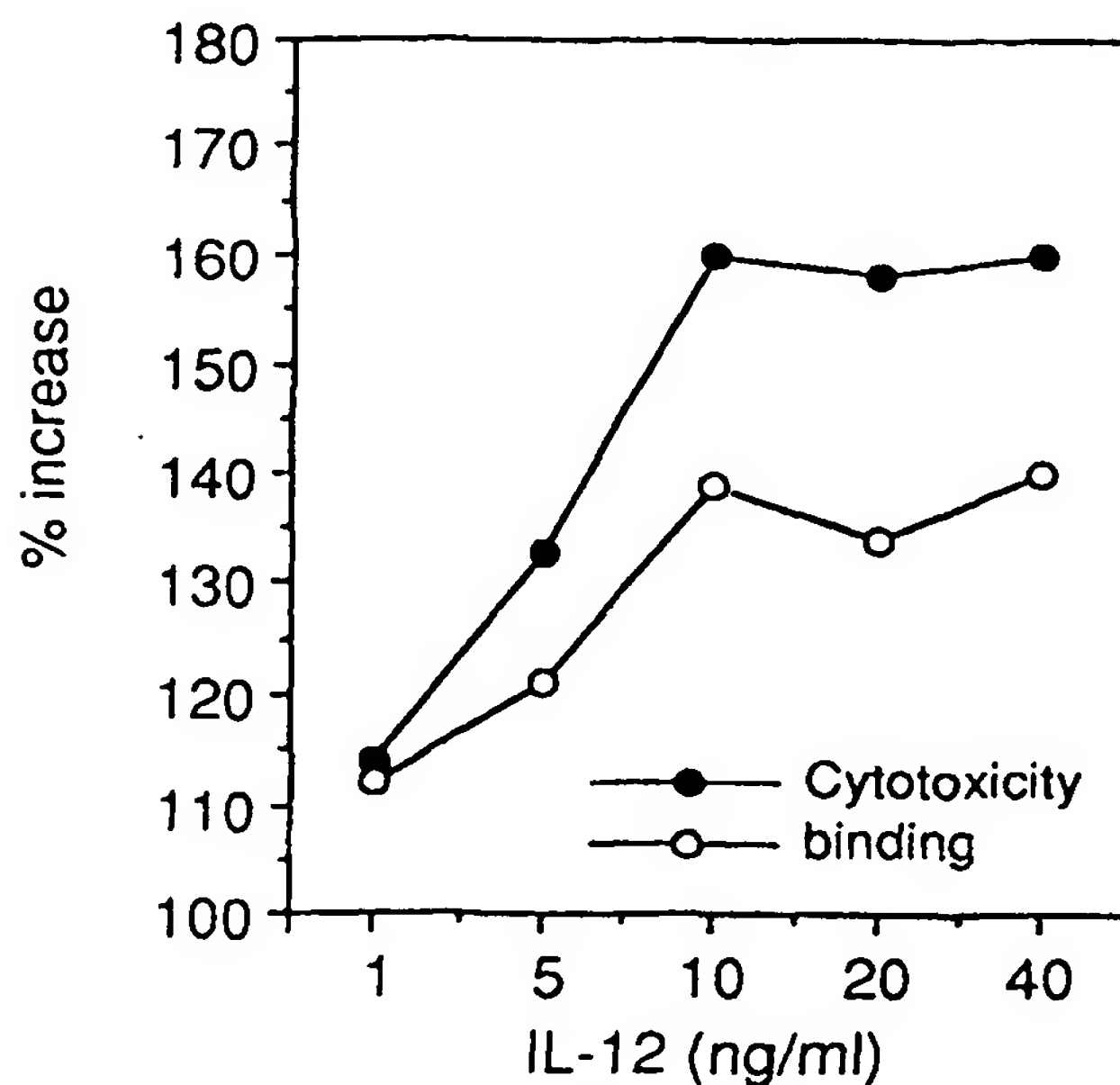
#### DISCUSSION

The chemotactic activity of NK cells is of particular interest for the immunobiology of these cells, as only few known cytokines were reported to have modulatory effects on their locomotor ability. IFN $\gamma$  and IL-2 have chemokinetic effects<sup>7,16</sup> and IL-8 has chemokinetic activity on long-term cultured IL-2-activated adherent NK cells, but the effect of this cytokine on resting NK cells was not tested.<sup>18</sup> The fact that IL-12 induces directional migration of freshly isolated

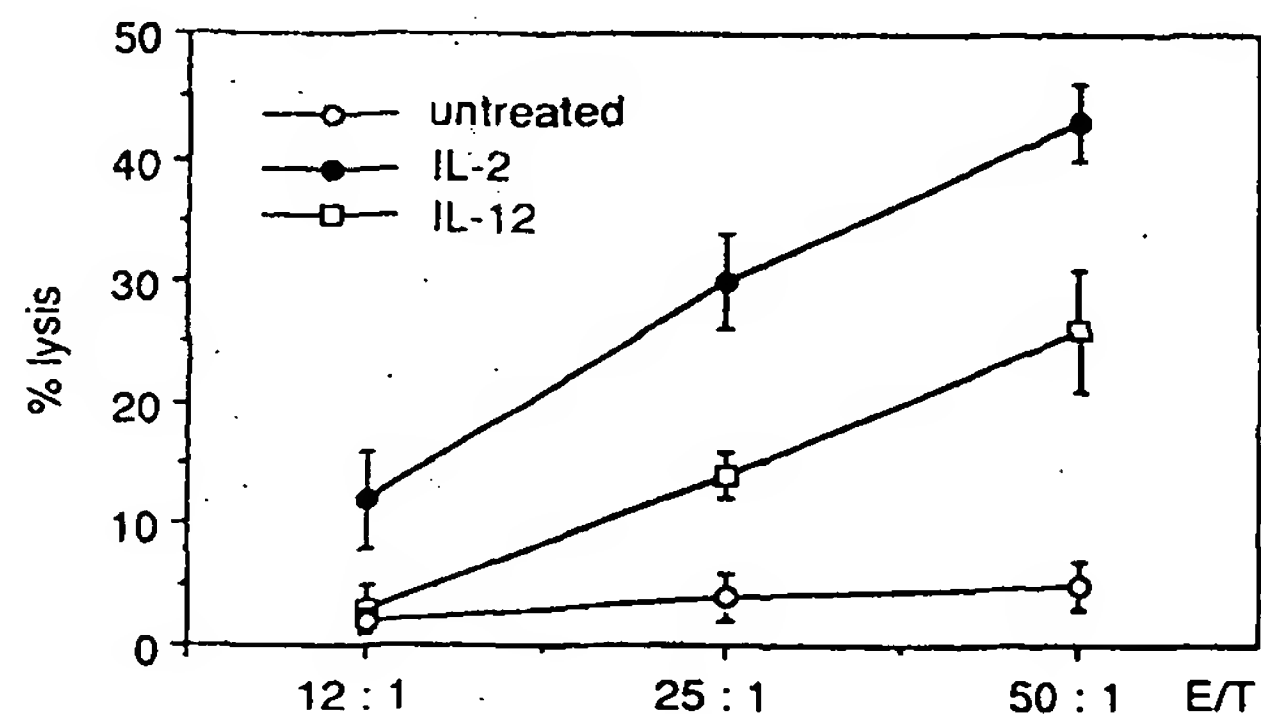
resting NK cells describes a new function for IL-12 and identifies a novel chemoattractant for this cell population.

In parallel experiments, it was found that CD3-activated T cells and PMN responded to a gradient of IL-12. The locomotor ability of T lymphocytes has recently received much attention after the reports that some members of the newly described chemokine family are active on T cells.<sup>35-38</sup> Macrophage inflammatory protein 1 $\alpha$  (MIP1 $\alpha$ ) was reported to preferentially chemoattract CD3-activated CD8 $^{+}$  T cells, while MIP1 $\beta$  preferentially attracted CD4 $^{+}$  T cells. RANTES was active on unstimulated CD4 $^{+}$ CD45RO $^{+}$  memory T cells and stimulated CD4 $^{+}$  and CD8 $^{+}$  lymphocytes. Finally, interferon-inducible protein 10 chemoattracted both stimulated CD4 $^{+}$  and CD8 $^{+}$  lymphocytes. In this study, IL-12 induced a chemotactic response only in CD3-activated, CD4 $^{+}$ , and CD8 $^{+}$  lymphocytes, indicating that the state of cell activation might be crucial for their recruitment in vivo.

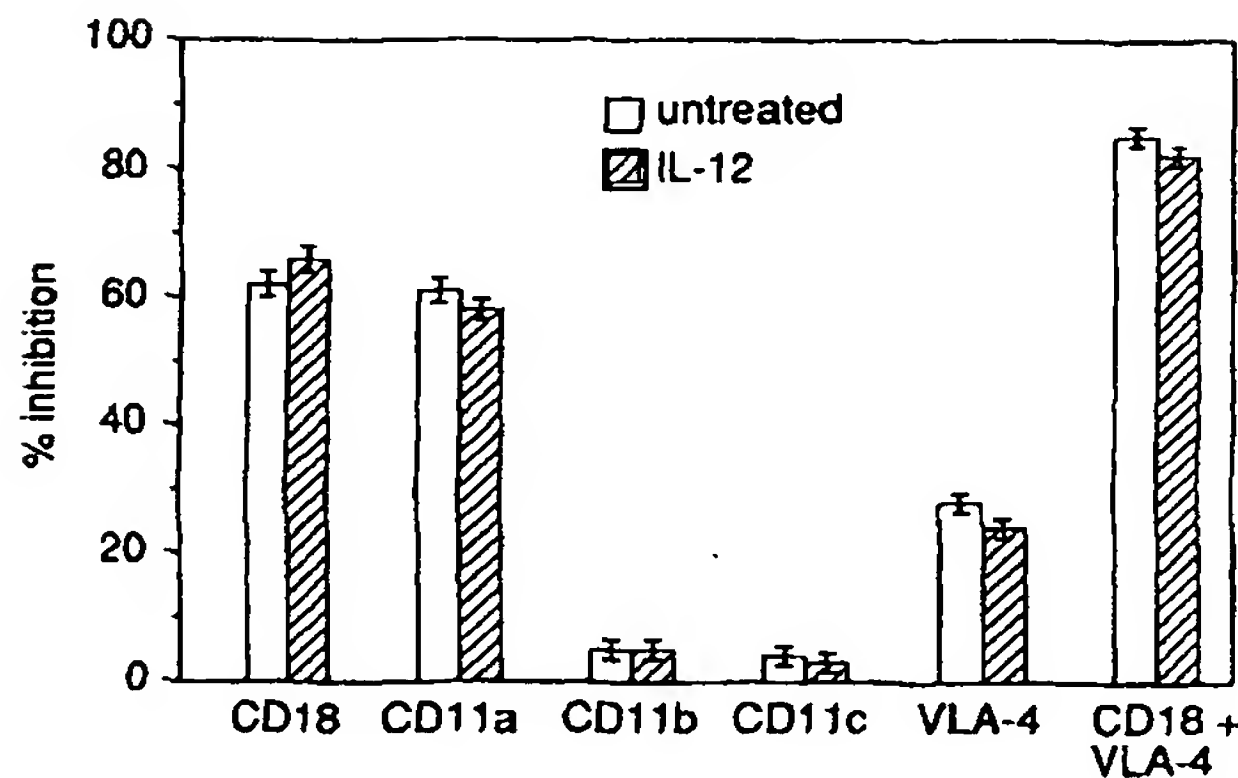
IL-12 was also chemotactic for PMN and chemotactic activity was abrogated by anti-IL-12 MoAb and by boiling the cytokine, thus excluding a nonspecific effect. Furthermore, the supernatant from mononuclear cells or purified PMN cultured for 2 hours with IL-12 (the length of the chemotaxis assay) was not active in inducing directional migration of freshly isolated PMN, suggesting that IL-12 is working directly, and not by inducing the secretion of a



**Fig 5.** Comparison of the dose response effect of IL-12 on NK-cell adhesion to resting EC and NK activity against the K562 cell line. Results are expressed as percent of increase assuming as 100% the activity of untreated cells (12% binding and 20 lytic units/ $10^5$  cells for cytotoxicity).



**Fig 6.** Effect of IL-12 on the cytotoxic activity of nonadherent lymphocytes against in vitro cultured human resting EC. Lymphocytes were incubated overnight with IL-12 (10 ng/mL), IL-2 (100 U/mL), or medium. Cytotoxicity was evaluated in a 4-hour assay using  $^{51}\text{Cr}$ -labeled EC as target.

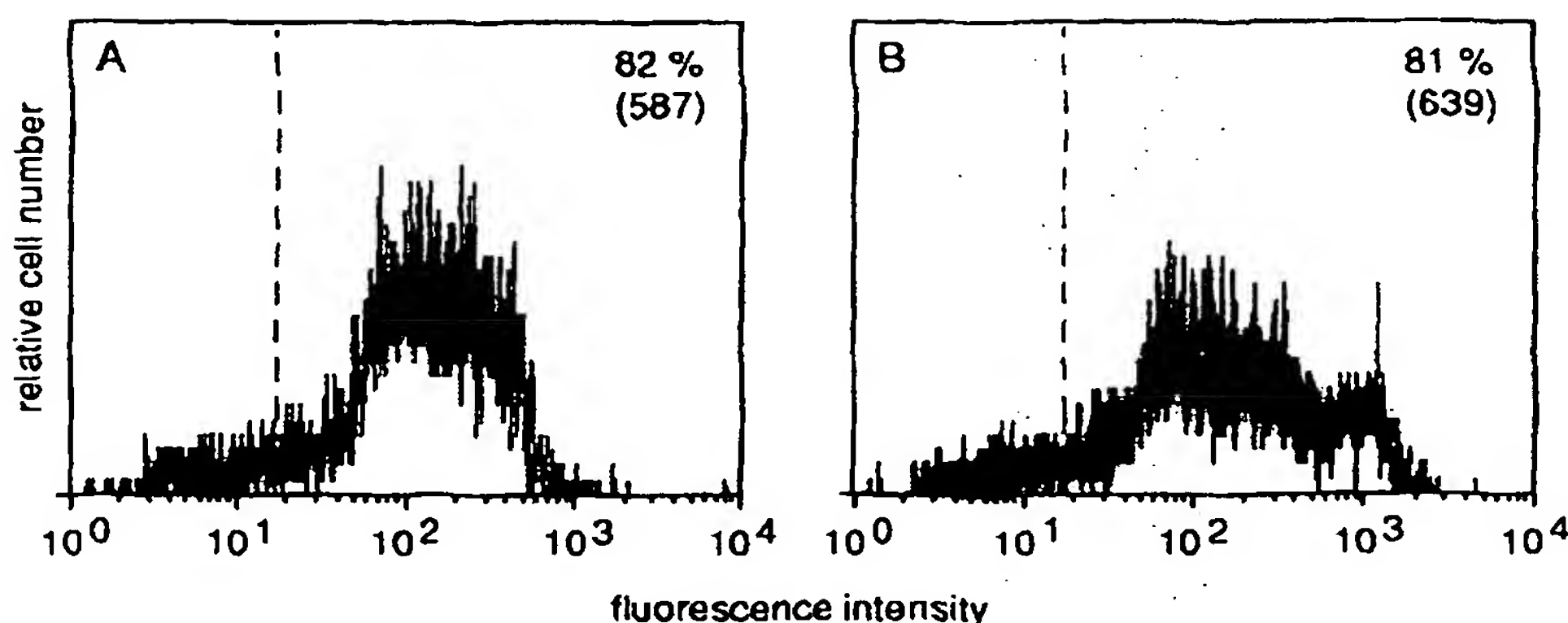


**Fig 7.** Adhesion molecules involved in IL-12-induced NK-cell binding to IL-1-activated EC. NK cells were treated overnight with IL-12 (10 ng/mL) or medium.  $^{51}\text{Cr}$ -labeled cells were incubated with optimal concentrations of MoAbs. Data are expressed as percent inhibition compared with control cells incubated with an irrelevant MoAb. Values are the mean  $\pm$  SD (four replicates) from one of three similar experiments.

second factor (data not shown). We failed to observe modulation of other functions of PMN by IL-12, including expression of mRNA levels for IL-6, IL-1 $\beta$ , IL-1 receptor antagonist (IRA), IL-1 receptor type I and type II, *c-fos*, *c-Jun*, and *c-myc*. The significance of this selective effect of IL-12 on mature myelocytes (monocytes were unaffected) remains to be defined.

We and others have reported that NK cells bind to EC and that this interaction is modulated by cytokines.<sup>12,19</sup> It was, therefore, of interest to study the action of IL-12 at this level. IL-12 augmented the number of NK cells that bound to resting EC (68%) and activated EC (33%). This effect of IL-12 showed several differences compared with the effect induced by IL-2, a prototypic NK-cell-stimulating cytokine: (1) IL-2 had a more marked effect compared with IL-12; (2)

the kinetics of IL-12 were slower, as it required overnight treatment, while a 1-hour exposure to IL-2 was sufficient to give a significant increase in NK-cell binding<sup>12</sup>; and (3) IL-2-augmented binding was inhibited by IL-4,<sup>20</sup> while the latter cytokine does not inhibit the effect of IL-12 (data not shown). IL-4 was also ineffective in inhibiting the IL-12-augmented NK-cell cytotoxicity.<sup>39</sup> It is of interest that IL-4-treated monocytes showed a decreased secretion of IL-12.<sup>40</sup> One of the mechanisms through which IL-12 might mediate effects on NK cells involves the adhesion molecules critical for the interaction between NK and EC. Robertson et al<sup>39</sup> have reported that IL-12 augments some adhesion molecules on NK cells, including CD11a, CD54, and CD56, and this could explain the increased cytotoxicity. Also, Rabinovich et al<sup>41</sup> have recently reported augmented expression of  $\beta 2$  integrin on IL-12-activated NK cells.<sup>41</sup> However, this increased expression was observed after a relatively long incubation with IL-12 (6 days' culture), while we detected NK-cell augmentation of tumor killing and EC binding as early as 24 hours after incubation. Moreover, the effect on chemotaxis was readily observed in a 2-hour assay. In this study, we tested the expression of  $\beta 2$  integrins by fluorescence-activated cell sorter (FACS) with the most widely used anti-CD11a/CD18 MoAb, and did not find a consistent modification of expression, despite increased binding to EC. However, we did find an increased expression of an activation epitope of LFA-1 recognized by MoAb NKI-L16.<sup>34</sup> Rabinovich et al<sup>41</sup> have reported that IL-12 did not induce NK-cell binding to fibronectin and laminin-coated plastic. Moreover,  $\alpha 4\beta 1$  integrin expression was not augmented after 6 days' incubation with IL-12.<sup>41</sup> These results further support that the IL-12-activated NK-cell binding to EC is not mediated by modulation of  $\alpha 4\beta 1$  integrin. Thus, some evidences concur with the conclusion that IL-12 augments the adhesive ability of NK cells by affecting the LFA-1 molecule. This effect is most likely due to activation of the molecule, rather than augmented expression.



**Fig 8.** Expression of CD11a (NKI-L16 epitope) by resting (A) and IL-12-activated (B) NK cells. Cells were labeled with anti-NKI-L16 ascites diluted 1:100 in RPMI 1640 + 10% FCS (complete medium) at room temperature for 30 minutes. All subsequent washings and labeling with FITC-conjugated second antibody were performed in complete medium. Cells were analyzed with a FACStar (Becton Dickinson). NKI-L16 expression by NK cells incubated with medium: 82% positive cells, mean channel fluorescence (MCF) 587; NK cells incubated with IL-12 (10 ng/mL, 72 hours), 81% positive cells, MCF 639.

In addition to acting as natural effectors, NK cells play an important role in the generation of T helper 1 (TH1)-versus TH2-type responses.<sup>42</sup> IL-12 has been shown to directly influence the generation of TH1 versus TH2 cells.<sup>43</sup> The results presented have suggested that monocyte-derived IL-12 may induce or amplify recruitment of NK cells, important determinants of TH1 versus TH2 responses.

# REFERENCES

1. Herberman RB, Ortaldo JR: NK cells: Their role in defence against disease. *Science* 214:24, 1981
2. Trinchieri G: Biology of natural killer cells. *Adv Immunol* 47:187, 1989
3. Robertson MJ, Ritz J: Biology and clinical relevance of human natural killer cells. *Blood* 76:2421, 1990
4. Holmberg LA, Springer KA, Ault KA: Natural killer activity in the peritoneal exudate of mice infected with listeria monocitogenes. *J Immunol* 127:1792, 1981
5. McIntyre KW, Welsh RM: Accumulation of natural killer and cytotoxic T large granular lymphocytes in the liver during virus infection. *J Exp Med* 164:1667, 1986
6. Nemlander A, Saksela E, Hayrj P: Are natural killer cells involved in allograft rejection? *Eur J Immunol* 13:348, 1983
7. Bottazzi B, Introna M, Allavena P, Villa A, Mantovani A: In vitro migration of human large granular lymphocytes. *J Immunol* 134:2316, 1985
8. Pohajdak B, Gomez J, Orr FW, Khalil N, Talgoy M, Greenberg AH: Chemotaxis of large granular lymphocytes. *J Immunol* 136:278, 1986
9. Pirelli A, Allavena P, Mantovani A: Activated adherent large granular lymphocytes/natural killer (LGL/NK) cells change their migratory behaviour. *Immunology* 65:651, 1988
10. Somersalo K, Tarkkanen J, Patarroyo M, Saksela E: Involvement of beta 2-integrins in the migration of human natural killer cells. *J Immunol* 149:590, 1992
11. Bender JR, Pardi R, Karasek MA, Engleman EG: Phenotypic and functional characterization of lymphocytes that bind human microvascular endothelial cells in vitro. Evidence for preferential binding of natural killer cells. *J Clin Invest* 79:1679, 1987
12. Allavena P, Paganin C, Martin Padura I, Peri G, Gaboli M, Dejana E, Marchisio PC, Mantovani A: Molecules and structures involved in the adhesion of natural killer cells to vascular endothelium. *J Exp Med* 173:439, 1991
13. Gismondi A, Morrone S, Humphries MJ, Piccoli M, Frati L, Santoni A: Human natural killer cells express VLA-4 and VLA-5, which mediate their adhesion to fibronectin. *J Immunol* 146:384, 1991
14. Somersalo K, Saksela E: Fibronectin facilitates the migration of human natural killer cells. *Eur J Immunol* 21:35, 1991
15. Bianchi G, Sironi M, Ghibaudi E, Selvaggini C, Elices M, Allavena P, Mantovani A: Migration of NK cells across endothelial cell monolayers. *J Immunol* 151:5134, 1993
16. Polentarutti N, Bottazzi B, Balotta C, Erroi A, Mantovani A: Modulation of the locomotory capacity of human large granular lymphocytes. *Cell Immunol* 101:204, 1986
17. Maghazachi AA: Tumor necrosis factor-alpha is chemokinetic for lymphokine-activated killer cells: Regulation by cyclic adenosine monophosphate. *J Leukoc Biol* 49:302, 1991
18. Sebok K, Woodside D, al Aoukaty A, Ho AD, Gluck S, Maghazachi AA: IL-8 induces the locomotion of human IL-2-activated natural killer cells. Involvement of a guanine nucleotide binding (Go) protein. *J Immunol* 150:1524, 1993
19. Aronson FR, Libby P, Brandon EP, Janicka MW, Mier JW: IL-2 rapidly induces natural killer cell adhesion to human endothelial cells. A potential mechanism for endothelial injury. *J Immunol* 141:158, 1988
20. Paganin C, Matteucci C, Censuales S, Mantovani A, Allavena P: IL-4 inhibits binding and cytotoxicity of NK cells to vascular endothelium. *Cytokine* 6:135, 1994
21. Kobayashi M, Fitz L, Ryan M, Hewick RM, Clark SC, Chan S, Loudon R, Sherman F, Perussia B, Trinchieri G: Identification and purification of natural killer cell stimulatory factor (NKSF), a cytokine with multiple biologic effects on human lymphocytes. *J Exp Med* 170:827, 1989
22. Wolf SF, Temple PA, Kobayashi M, Young D, Dicig M, Lowe L, Dzialo R, Fitz L, Ferenz C, Hewick RM, Kelleher K, Herrmann SH, Clark SC, Azzoui L, Chan SH, Trinchieri G, Perussia B: Cloning of cDNA for natural killer cell stimulatory factor, a heterodimeric cytokine with multiple biologic effects on T and natural killer cells. *J Immunol* 146:3074, 1991
23. Gubler U, Chua AO, Schoenhaut DS, Dwyer CM, McComas W, Motyka R, Nabavi N, Wolitzky AG, Quinn PM, Familletti PC, Gately MK: Coexpression of two distinct genes is required to generate secreted bioactive cytotoxic lymphocyte maturation factor. *Proc Natl Acad Sci USA* 88:4143, 1991
24. Locksley RM: Interleukin 12 in host defense against microbial pathogens. *Proc Natl Acad Sci USA* 90:5879, 1993
25. Chan SH, Perussia B, Gupta JW, Kobayashi M, Pospisil M, Young HA, Wolf SF, Young D, Clark SC, Trinchieri G: Induction of interferon gamma production by natural killer cell stimulatory factor: Characterization of the responder cells and synergy with other inducers. *J Exp Med* 173:869, 1991
26. Gately MK, Desai BB, Wolitzky AG, Quinn PM, Dwyer CM, Podlaski FJ, Familletti PC, Sinigaglia F, Chizzonite R, Gubler U, Stern AS: Regulation of human lymphocyte proliferation by a heterodimeric cytokine, IL-12 (cytotoxic lymphocyte maturation factor). *J Immunol* 147:874, 1991
27. Sozzani S, Luini W, Molino M, Jilek P, Bottazzi B, Cerletti C, Matsushima K, Mantovani A: The signal transduction pathway involved in the migration induced by a monocyte chemotactic cytokine. *J Immunol* 147:2215, 1991
28. Colotta F, Peri G, Villa A, Mantovani A: Rapid killing of actinomycin D-treated tumor cells by human mononuclear cells. I. Effectors belong to the monocyte-macrophage lineage. *J Immunol* 132:936, 1984
29. Falk W, Goodwin Jr RH, Leonard EJ: A 48-well microchemotaxis assembly for rapid and accurate measurement of leukocyte migration. *J Immunol Methods* 33:239, 1980
30. Zigmond SH, Hirsch JG: Leukocyte locomotion and chemotaxis: New methods for evaluation and demonstration of a cell-derived chemotactic factor. *J Exp Med* 137:387, 1973
31. Sironi M, Breviario F, Proserpio P, Biondi A, Vecchi A, Van Damme J, Dejana E, Mantovani A: IL-1 stimulates IL-6 production in endothelial cells. *J Immunol* 142:549, 1989
32. Damle NK, Doyle LV, Bender JR, Bradley EC: Interleukin 2-activated human lymphocytes exhibit enhanced adhesion to normal vascular endothelial cells and cause their lysis. *J Immunol* 138:1779, 1987
33. Mier JW, Brandon EP, Libby P, Janicka MW, Aronson FR: Activated endothelial cells resist lymphokine-activated killer cell-mediated injury. Possible role of induced cytokines in limiting capillary leak during IL-2 therapy. *J Immunol* 143:2407, 1989
34. Figdor CG, va Kooyk I, Keizer GD: On the mode of action of LFA-1. *Immunol Today* 11:277, 1990
35. Tanaka Y, Adams DH, Hubscher S, Hirano H, Siebenlist U, Shaw S: T-cell adhesion induced by proteoglycan-immobilized cytokine MIP-1 beta. *Nature* 361:79, 1993



36. Taub DD, Conlon K, Lloyd AR, Oppenheim JJ, Kelvin DJ: Preferential migration of activated CD4+ and CD8+ T cells in response to MIP-1 alpha and MIP-1 beta. *Science* 260:355, 1993
37. Schall TJ, Bacon K, Camp RD, Kaspari JW, Goeddel DV: Human macrophage inflammatory protein alpha (MIP-1 alpha) and MIP-1 beta chemokines attract distinct populations of lymphocytes. *J Exp Med* 177:1821, 1993
38. Taub DD, Lloyd AR, Conlon K, Wang JM, Ortaldo JR, Harada A, Matsushima K, Kelvin DJ, Oppenheim JJ: Recombinant human interferon-inducible protein 10 is a chemoattractant for human monocytes and T lymphocytes and promotes T cell adhesion to endothelial cells. *J Exp Med* 177:1809, 1993
39. Robertson MJ, Soiffer RJ, Wolf SF, Manley TJ, Donahue C, Young D, Herrmann SH, Ritz J: Response of human natural killer (NK) cells to NK cell stimulatory factor (NKSF): Cytolytic activity and proliferation of NK cells are differentially regulated by NKSF. *J Exp Med* 175:779, 1992
40. Trinchieri G: Interleukin-12 and its role in the generation of Th1 cells. *Immunol Today* 14:335, 1993
41. Rabinovich H, Herberman RB, Whiteside T: Differential effects of IL-12 and IL-2 in expression and function of cellular adhesion molecules on purified NK cells. *Cell Immunol* 152:481, 1993
42. Scharon TM, Scott P: Natural killer cells are a source of Interferon  $\gamma$  that drives differentiation of CD4+ T cell subset and induces early resistance to *Leishmania major* in mice. *J Exp Med* 178:567, 1993
43. Manetti R, Parronchi P, Giudizi MG, Piccinini MP, Maggi E, Trinchieri G, Romagnani S: Natural killer cell stimulatory factor (interleukin 12 [IL-12]) induces T helper type 1 (Th1)-specific immune responses and inhibits the development of IL-4-producing Th cells. *J Exp Med* 177:1199, 1993



## Cloning and Characterization of a Specific Interleukin (IL)-13 Binding Protein Structurally Related to the IL-5 Receptor $\alpha$ Chain\*

(Received for publication, February 5, 1996, and in revised form, April 5, 1996)

Daniel Caput, Patrick Laurent, Mourad Kaghad, Jean-Michel Lelias, Sylvie Lefort, Natalio Vita, and Pascual Ferrara†

From Sanofi Recherche, BP 137, 31676 Labège Cedex, France

Interleukin-13 (IL-13) is a cytokine secreted by activated T lymphocytes that shares many, but not all, biological activities with IL-4. These overlapping activities are probably due to the existence of common receptor components. Two proteins have been described as constituents of the IL-4 receptor, a ~140-kDa glycoprotein (IL-4R) and the  $\gamma$  chain ( $\gamma$ c) of the IL-2 receptor, but neither of these proteins binds IL-13. We have cloned a cDNA encoding an IL-13 binding protein (IL-13R) from the Caki-1 human renal carcinoma cell line. The cloned cDNA encodes a 380-amino acid protein with two consensus patterns characteristic of the hematopoietic cytokine receptor family and a short cytoplasmic tail. The IL-13R shows homology with the IL-5 receptor, and to a lesser extent, with the prolactin receptor. COS-7 cells transfected with the IL-13R cDNA bind IL-13 with high affinity but do not bind IL-4. COS-7 cells co-transfected with the cloned IL-13R cDNA and IL-4R cDNA resulted in the reconstitution of a small number of receptors that recognized both IL-4 and IL-13. Reverse transcription-polymerase chain reaction analysis detected the receptor transcript only in cell lines known to bind IL-13.

IL-13<sup>1</sup> is a cytokine secreted by activated T lymphocytes that regulates inflammatory and immune responses (1, 2). It shares several biological activities with IL-4, another T-cell-derived cytokine, in a variety of cell types such as B cells, monocytes, fibroblasts, and endothelial cells (for review, see Ref. 3). But contrary to IL-4, IL-13 does not regulate T-lymphocyte function (4). Recent studies have shown that IL-13 competitively inhibits IL-4 binding to the receptor present on some cell lines, whereas on other cell lines, IL-4 binding cannot be displaced by IL-13 (5, 6). Similarly, IL-13 binding was either completely or partially displaced by IL-4, depending on the cell line used as source of receptors (6, 7). It has also been shown that a mutated IL-4 (8) or anti-IL-4R antibodies that block the biological activity of IL-4 also antagonize IL-13 activity (5, 9–11) and that the signaling events induced by IL-4 and IL-13 are similar in the cells that respond to both cytokines (10, 12, 13). Together, these results suggest that IL-4 and IL-13 receptors are not the same but that they share receptor components.

Two proteins have been described as components of the high affinity IL-4R complex: IL-4R, a glycoprotein of 140 kDa,

which, when expressed in COS cells, binds IL-4 with a  $K_d$  of 50–100 pM (14, 15); and the  $\gamma$  chain of the IL-2 receptor (16, 17), which, when associated with the IL-4R, results in a 2–3-fold increase in affinity for IL-4 and also participates in some of the IL-4-mediated signal transduction events (17). Neither of these two proteins binds IL-13, nor does the complex of the two proteins (6). Thus, other protein(s), part of the IL-4 complex, are necessary for the recognition of IL-13.

Recently, it has been discovered that some human renal carcinoma cells express, in addition to receptors shared by IL-4 and IL-13, a large excess of specific IL-13 receptors (7). In view of the importance for molecular cloning of having a cell line expressing a large number of IL-13 receptors, we screened a panel of human carcinoma cell lines for IL-13 binding. We found one, the Caki-1 cell line, that expressed a high number of binding sites for IL-13, of which only a small fraction were shared with IL-4. We describe here the characterization of the IL-13 binding protein (IL-13R) present in these cells, its cloning, and the characterization of the recombinant protein expressed in COS-7 cells.

### MATERIALS AND METHODS

**Growth Factors, Antibodies, and Cells**—Recombinant IL-13 was produced and purified in our laboratory as described previously (2). Human IL-4 was obtained from Tebu (Le Perray en Yvelines, France). The anti-IL-4 receptor antibodies X2/45 and CDW124 and the IL-4 antagonist hIL-4.Y124D were a generous gift of Dr. Sebald (Würzburg, Germany). Caki-1 cells (ATCC HTB 46) were cultured in RPMI 1640 medium, 20% fetal calf serum, 2 mM glutamine, and penicillin/streptomycin (100 units/ml), at 37 °C in a humidified atmosphere containing 5% CO<sub>2</sub>.

**Binding and Biological Activity Assays**—Binding and cross-linking experiments were performed as described with <sup>125</sup>I-labeled [Phe<sup>43</sup>] IL-13-Gly-Tyr-Gly-Tyr (6) or <sup>125</sup>I-IL-4. For the induction of IL-6 secretion, Caki-1 cells were plated in 24-well plates at a density of 5 × 10<sup>4</sup> cells/well, and after 3 days of culture, confluent monolayers were washed three times with Dulbecco's modified Eagle's medium without fetal calf serum. Stimulation of Caki-1 cells was performed with 30 ng/ml of IL-4 or IL-13 in the absence or in the presence of hIL-4.Y124D or an anti-gp140 monoclonal antibody. The level of IL-6 released into the culture media after 24 h incubation was measured by enzyme-linked immunosorbent assay (Innotest, Besançon, France).

**cDNA Library Constructions, Isolation of cDNAs, and Sequence Analysis**—Total RNA was extracted from Caki-1 cells or tissues as described (2). Poly(A) RNA was isolated from total RNA with oligo(dT)<sub>25</sub> magnetic beads (Dynal, Inc.). The Caki-1 size fractionated cDNA library containing 2 × 10<sup>5</sup> clones was constructed using the primer-adaptor procedure (18) and the vector pSE1 (19). The expression cloning strategy used was reported previously (20). The human cDNA sequence for the IL-13R has been submitted to the GenBank™/EMBL Data Bank (accession number X95302).

**Cell Transfection and Receptor Characterization**—For the functional characterization of the recombinant receptor, COS-7 cells were transfected in 25-cm<sup>2</sup> plates using 0.6 µg of different plasmids. After 24 h COS-7 cell monolayers were trypsinized and plated at 8 × 10<sup>4</sup> cells per well in 12-well plates. Three days later, binding and cross-linking experiments were performed as described with <sup>125</sup>I-labeled [Phe<sup>43</sup>] IL-13-Gly-Tyr-Gly-Tyr (6) or <sup>125</sup>I-IL-4.

\* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) X95302.

† To whom correspondence should be addressed. Tel: (33) 61 00 40 00; Fax: (33) 61 00 40 01; E-mail: pascual.ferrara@t1s1.elfsanofi.fr.

<sup>1</sup> The abbreviations used are: IL, interleukin; PCR, polymerase chain reaction;  $\gamma$ c,  $\gamma$  chain; R, receptor.

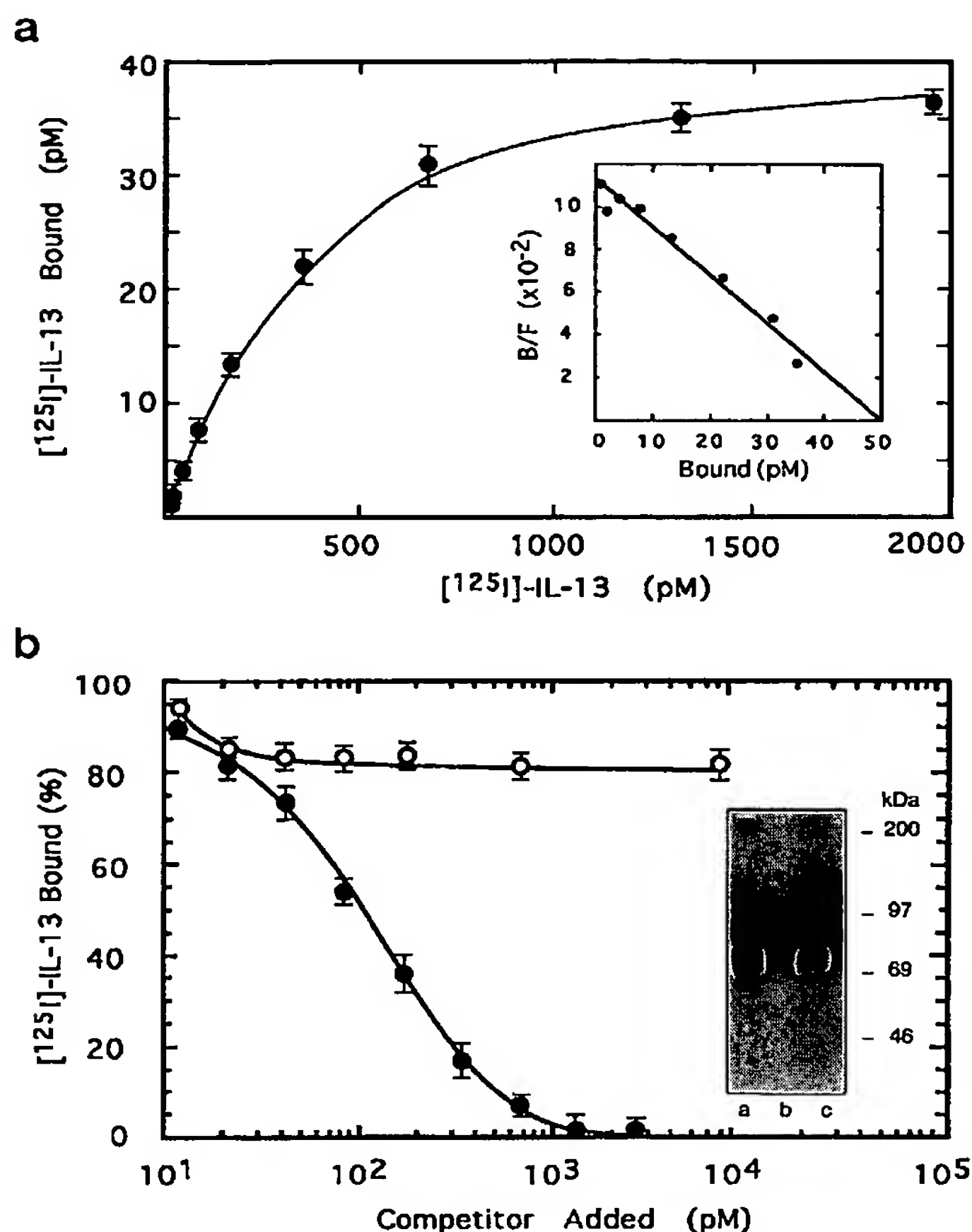


FIG. 1. Characterization of the IL-13R present in Caki-1 cells. *a*, Scatchard analysis (*inset*) of the  $^{125}\text{I}$ -IL-13 saturation curve, which indicated the presence of  $\sim 76,000$  sites/cell with a  $K_d \approx 440$  pM; *b*, binding of  $^{125}\text{I}$ -IL-13 in the presence of increasing concentrations of competing IL-13 ( $\bullet$ ) and IL-4 ( $\circ$ ). Bars, S.D. ( $n = 3$ ); *Inset*, cross-linking experiments using  $^{125}\text{I}$ -IL-13 in the absence (*lane a*) and in the presence of a 100-fold excess of IL-13 (*lane b*) or IL-4 (*lane c*).

**Tissue Distribution**—cDNAs were generated from RNA samples with reverse transcriptase and a mixture of labeled dNTP of known specific activity for quantification. Then, for each sample, 10 ng of cDNA was submitted to PCR using a sense primer corresponding to the sequence +52 to +71 and an antisense primer corresponding to +489 to +470 (numbering is based on the cDNA sequence shown in Fig. 3a). PCR-amplified products were hybridized with a probe complementary to sequence +445 to +461 of the cDNA.

## RESULTS

**Characterization of IL-4 and IL-13 Binding and Activity in Caki-1 Cells**—The Caki-1 cell line, among several human cell lines tested, expressed the highest number of binding sites for IL-13. Saturation experiments with labeled IL-13 on Caki-1 cells showed the presence of one class of binding sites with a  $K_d$  of  $446 \pm 50$  pM and  $7.2 \times 10^4$  receptors/cell (Fig. 1a). In competition experiments, unlabeled IL-13 totally displaced the labeled IL-13 in a dose-dependent manner, whereas IL-4 displaced with high affinity  $\sim 10\%$  of the labeled IL-13. Higher concentrations of IL-4 ( $>100$  nM) did not displace the remaining 90% of the IL-13 binding (Fig. 1b). These results are in line with the existence of two sites with similar affinity for IL-13, of which only one ( $\sim 10\%$  of the total binding) recognizes IL-4. Affinity cross-linking of IL-13 showed a  $\sim 70$ -kDa complex consistent with the complex observed in similar cross-linking experiments with IL-13 on several cell types (6). The labeled IL-13 was completely displaced from the complex by IL-13 but not by IL-4, which is in agreement with the competition

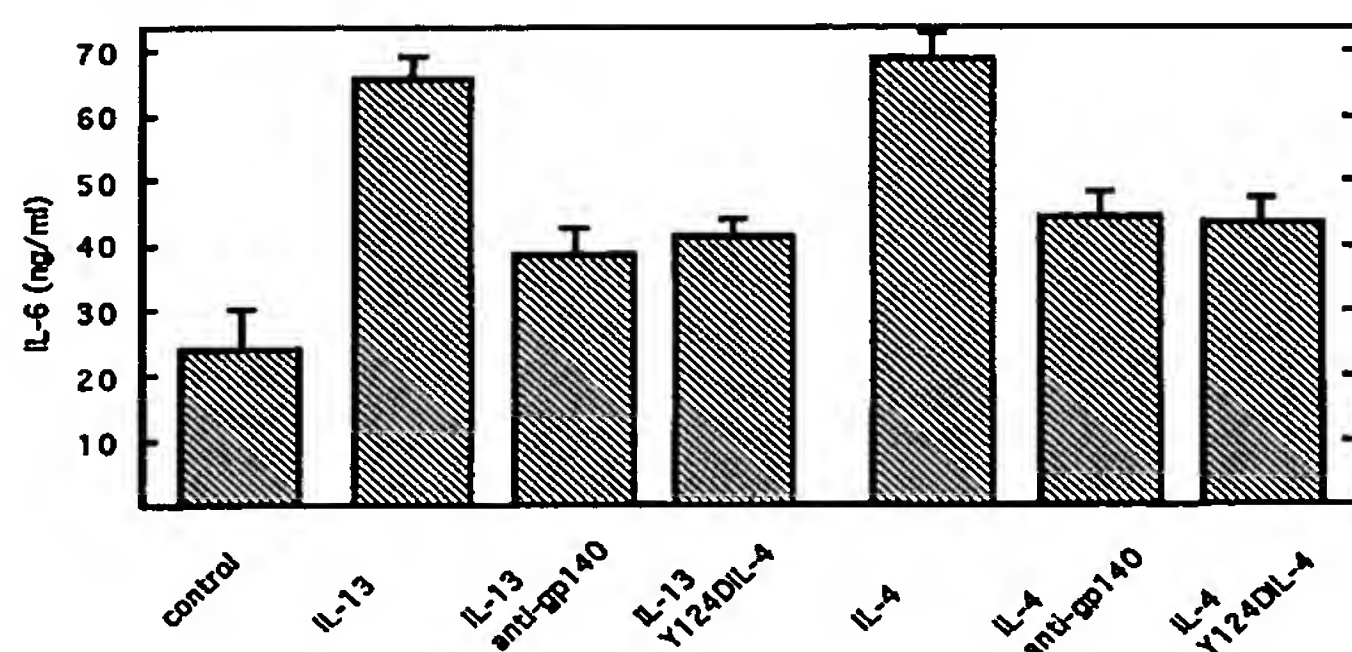


FIG. 2. Inhibition of the secretion of IL-6 induced by IL-13 and IL-4 in the presence of the monoclonal antibody anti-IL-4R and the IL-4 antagonist hIL-4.Y124D. Stimulation of Caki-1 cells was performed with 30 ng/ml of IL-4 or IL-13 in the absence or in the presence of hIL-4.Y124D or an anti-gp140 monoclonal antibody. The level of IL-6 released into the culture media after 24 h incubation was determined as described under "Materials and Methods." Each value represents the mean  $\pm$  S.E. of triplicate determinations.

experiments (Fig. 1b).

We also analyzed the IL-4- or IL-13-induced IL-6 secretion from Caki-1 cells. Both cytokines induced the secretion of similar levels of IL-6, and this secretion was inhibited by antibodies against the IL-4R $\alpha$  chain and by the IL-4 antagonist hIL-4.Y124D (Fig. 2), suggesting that only the shared receptors in Caki-1 cells are responsible for the induction of IL-6 secretion. Similar results were observed when the IL-4- and IL-13-induced phosphorylation of the IRS1/4PS was analyzed in the presence or the absence of the anti-IL-4R antibodies and the IL-4 antagonist (results not shown).

**Cloning of the IL-13 Binding Protein**—We constructed a cDNA library containing  $2 \times 10^5$  recombinant clones from Caki-1 cells. The library was divided into pools of 1000 cDNAs, and plasmid DNA from each pool was introduced into COS-7 cells. Binding of labeled IL-13 to transfected COS-7 cells was used to identify pools of clones encoding an IL-13R. Positive pools were partitioned and rescreened until a single clone was identified that directed synthesis of a cell surface protein capable of binding IL-13. Two independent IL-13 receptor cDNAs of identical sequence were finally isolated. The cDNA is 1299 bases long, excluding the poly(A) tract, and has a short 3'-untranslated region of 103 bases. A canonical AATAAA polyadenylation signal is found at the predicted location (Fig. 3a). The open reading frame between nucleotides 53 and 1192 defines a polypeptide of 380 amino acids. The sequence codes for a membrane protein with a putative signal peptide of 26 amino acids, a single membrane-spanning domain, and a short cytoplasmic tail. Four sites for potential N-linked glycosylation are located in the extracellular region. Importantly, two consensus patterns considered as signatures of the hematopoietic cytokine receptor family (for review, see Ref. 21) are also found, as are four conserved cysteines in the amino-terminal half of the extracellular domain and the WSXWS motif located in the COOH-terminal region of the extracellular domain (Fig. 3b). Alignment studies reveal homologies with the human IL-5R $\alpha$  chain (51% similarity and 27% identity; Fig. 3c) and to a lesser extent with the prolactin receptor (not shown).

**Distribution of the mRNA for the IL-13R**—A  $\sim 1.4$ -kilobase IL-13R transcript was detected when the cDNA for the IL-13R was labeled with  $^{32}\text{P}$  and used to probe a Northern blot containing polyadenylated RNA from Caki-1 cells (not shown). Surprisingly, in Caki-1 cells, similar amounts of IL-13R and IL-4R mRNAs are detected by Northern analysis, although a large excess of IL-13R is expressed. This observation suggests a higher translatability of this mRNA versus IL-4R transcripts and may explain the absence of IL-13R mRNA detection in cell



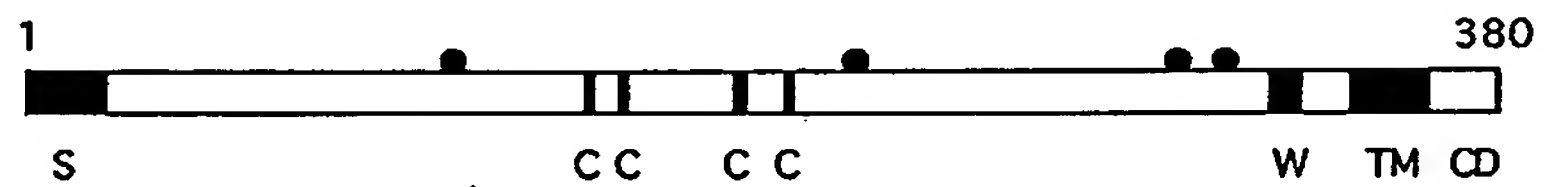
a

```

1  GGTGCCTGTCGGCGGGGAGAGAGGCAATATCAAGGTTTAAATCTCGGAGAAATGGCTTT 60
    M A F
61  CGTTTGCTTGGCTATCGGATGCTTATATACCTTTCTGATAAGCACAACTTTGGCTGTAC 120
4   V C L A I G C L Y T F L I S T T F G C T 23
121 TTCATCTTCAGACACCGAGATAAAGTTAAACCTCCTCAGGATTTGAGATAGTGGATCC 180
24  S S S D T E I K V N P P Q D F E I V D P 43
181 CGGATACTTAGGTTATCTCTATTTGCAATGGCAACCCCACTGTCTCTGGATCATTTTAA 240
44  G Y L G Y L Y L Q W Q P P L S L D H F K 63
241 GGAATGCACAGTGAATATGAATAAAATACCGAAACATTGGTAGTGAACATGGAAGAC 300
64  E C T V E Y E L K Y R N I G S E T W K T 83
301 CATCATTACTAAGAATCTACATTACAAAGATGGGTTTGTATCTTAACAAGGGCATTGAAGC 360
84  I I T K N L H Y K D G F D L N K G I E A 103
361 GAAGATACACACGCTTTTACCATGGCAATGCACAAATGGATCAGAAGTTCAAAGTTCTG 420
104 K I H T L L P W Q C T N G S E V Q S S W 123
421 GGCAGAACTACTTATTGGATATCACCACAAGGAATCCAGAACTAAAGTTCCAGGATAT 480
124 A E T T Y W I S P Q G I P E T K V D M 143
481 GGATTGCGTATATTACAATTGGCAATATTTACTCTGTCTTGGAAACCTGGCATAGGTGT 540
144 D C V Y Y N W Q Y L L C S W K P G I G V 163
541 ACTTCTTGATACCAATTACAATTGTTTACTGGTATGAGGGCTGGATCATGCATTACA 600
164 L L D T N Y N L F Y W Y E G L D H A L Q 183
601 GTGTGTGATTACATCAAGGCTGATGGACAAATATAGGATGCAGATTTCCCTATTGGA 660
184 C V D Y I K A D G Q N I G C R F P Y L E 203
661 GGCATCAGACTATAAAGATTCTATATTTGTGTTAATGGATCATCAGAGAACAAGCCTAT 720
204 A S D Y K D F Y I C V N G S S E N K P I 223
721 CAGATCCAGTTATTTCACTTTTCAGCTTCAAATATAGTTAAACCTTTGCCGCCAGTCTA 780
224 R S S Y F T F Q L Q N I V K P L P Y Y 243
781 TCTTACTTTTACTCGGGAGAGTTTATGTGAATTAAGCTGAATGGAGCATACCTTTGGG 840
244 L T F T R E S S C E I K L K W S I P L G 263
841 ACCTATCCAGCAAGGTGTTTGTATGAAATGAGATCAGAGAAGATGATACCTT 900
264 P I P A R C F D Y E I E I R E D D T T L 283
901 GGTGACTGCTACAGTTGAAAATGAAACATACACCTTGAAAACACAAATGAAACCCGACA 960
284 V T A T V E N E T Y T L K T T N E T R Q 303
961 ATTATGCTTTGTAGTAAGAAGCAAAGTGAATATTTATGCTCAGATGACGGAATTTGGAG 1020
304 L C F V V R S K V N I Y C S D D G I W S 323
1021 TGAGTGGAGTGATAAACAATGCTGGGAAGGTGAAGACCTATCGAAGAAACTTTGTACG 1080
324 E W S D K Q C N E G E D L S K K T L R 343
1081 TTTCTGGCTACCATTTGGTTTCACTTAATATTAGTTATATTGTAAACCGGTCTGCTTTT 1140
344 F N L P F G F I L I L V I F V T G L L L 363
1141 GCGTAAGCCAAACACCTACCCAAAAATGATCCAGAATTTTCTGTGATACATGAAGACT 1200
364 R K P N T Y P K M I P E F F C D T 380
1201 TTCCATATCAAGAGACATGGTATTGACTCAACAGTTCCAGTCATGGCCAAATGTTCAAT 1260
1261 ATGAGTCTCAATAAAGTGAATTTTCTTGGCAATGTTG 1298

```

b



c

```

IL13R  MAFVCLAIGCLYTFLLSTTFGCTSSSDTEIKVNPPQDFEIVDPGYLGPLY 50
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
IL5R   ..MIIVAHVLLILLGATEILQADLLPDEKISLLPPVNFITKVTLAQVL 47

IL13R  LQWQPFSLDHFKECTVEYELKYRNIGSETWKTITKNLHYKDGFDLNKG 100
      ||| | | | | | | | | | | | | | | | | | | | | | | | | |
IL5R   LQWKPNPDQEQ.RNVNLEYQVKINAPKEDDYETRITES...KCVTILHKG 93

IL13R  IEAKIHTLLFPWQCTNGSEVQSSWAETTYWISPOGIPETKVQDMICV.... 146
      | | | | | | | | | | | | | | | | | | | | | | | | | |
IL5R   FSASVRTILQ...NDHSLASSWASAE.LHAPPGSPGTSIVNLICTTNTT 139

IL13R  ..YYNWQ.....YLLCSWKPGIGVLLDTNYNLFYWYEGLDHALQCVDYIK 189
      | | | | | | | | | | | | | | | | | | | | | | | | | |
IL5R   EDNYSRLRSYQVSLHCTFWLVGTDAPEDTQYFLYRYGSWTE..EQEYSK 187

IL13R  AD.GQNIGCRFP..YLEASDYKDFYICVNGSSSENKPIRSSYFTFQLQNI 236
      ||| | | | | | | | | | | | | | | | | | | | | | | |
IL5R   DTLGRNIACWFFRTFILSKGRDWLSVLVNGSSKHSAIRPFDQLFALHAID 237

IL13R  KPLPPVYLTFTRESSCEIKLKWSIPLGPIPARCFDYIEIREDDTTLVTA 286
      || | | | | | | | | | | | | | | | | | | | | | | | |
IL5R   QINPPLNVTAIEIEGT.RLSIQWEKPVSAFPIHCFDYEVKIHNTNRYLQI 286

IL13R  TVENETYTLKTTNETRQLCFVVRSKVNIYCSDDGIIWSEWSDKQCWEGEDL 336
      | | | | | | | | | | | | | | | | | | | | | | | | | |
IL5R   EKLMTNAFISIIDDLISKYDVQVRAAVSSMCREAGLWSEWSQ.PIYVGND 335

IL13R  SKKTLLRFWLPFGFILILVIFVTGLLLRKPNTYPKMIP.....EF 376
      | | | | | | | | | | | | | | | | | | | | | | | | | |
IL5R   HKPLREWFVIVIMATICFILLILSLICKICHLWIKLFPPIPAPKSNIKDL 385

IL13R  FCDT..... 380
      | | | | | | | | | | | | | | | | | | | | | | | | | |
IL5R   FVTNIEKAGSSETEIEVICYIEKPGVETLEDVSF 420

```

FIG. 3. Nucleotide sequence of the IL-13R cDNA and comparison of the IL-13R and IL-5R protein sequences. *a*, nucleotide sequence of the IL-13R cDNA. The human IL-13R cDNA sequence has been submitted to the GenBank™/EMBL Data Bank (accession number X95302). The amino acids corresponding to the predicted transmembrane domain are underlined. Potential *N*-glycosylation sites (Asn-X-Ser/Thr) are underscored. *b*, schematic representation of the protein. *S*, predicted signal peptide; *C*, conserved cysteines in the hematopoietic cytokine receptor family; *W*, the WSXWS motif; *TM*, transmembrane domain; *CD*, cytoplasmic domain. The *black circles* indicate the potential *N*-glycosylation sites. *c*, amino acid alignment of the IL-13R and IL-5R sequences. Human IL-13 and IL-5 receptor protein sequences are aligned as described (27). Cysteine residues and the WSXWS motif characteristic of this family of receptors are *boxed*.

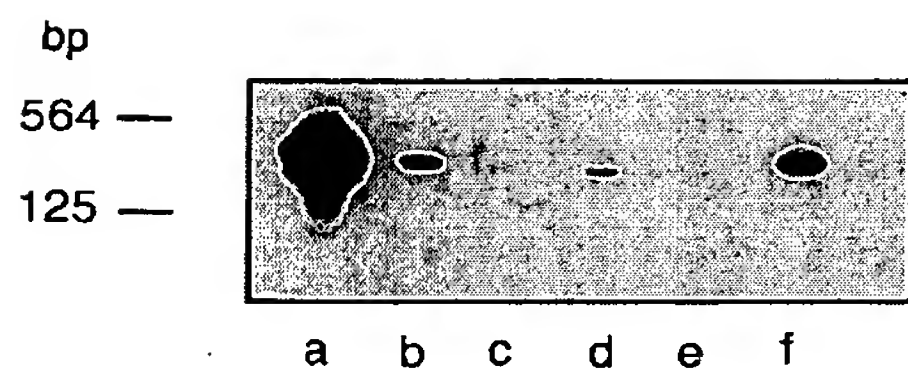


FIG. 4. Expression patterns of the IL-13 receptor mRNA. RNA was prepared from the following cells: Caki-1 (lane a), A431 (lane b), TF-1 (lane c), U937 (lane d), Jurkat (lane e), and IM9 (lane f). RNA samples were copied with reverse transcriptase and submitted to PCR, and amplified products were hybridized with a probe complementary to sequence +445 to +461 of the cDNA as described under "Materials and Methods." Size markers are indicated at the left of the figure. The estimated number of receptors/cell for Caki-1, A431, TF-1, U937, Jurkat, and IM9 cells are  $\approx 60000$ , 200, 250, 420,  $<25$ , and 400 receptors/cell, respectively (6).

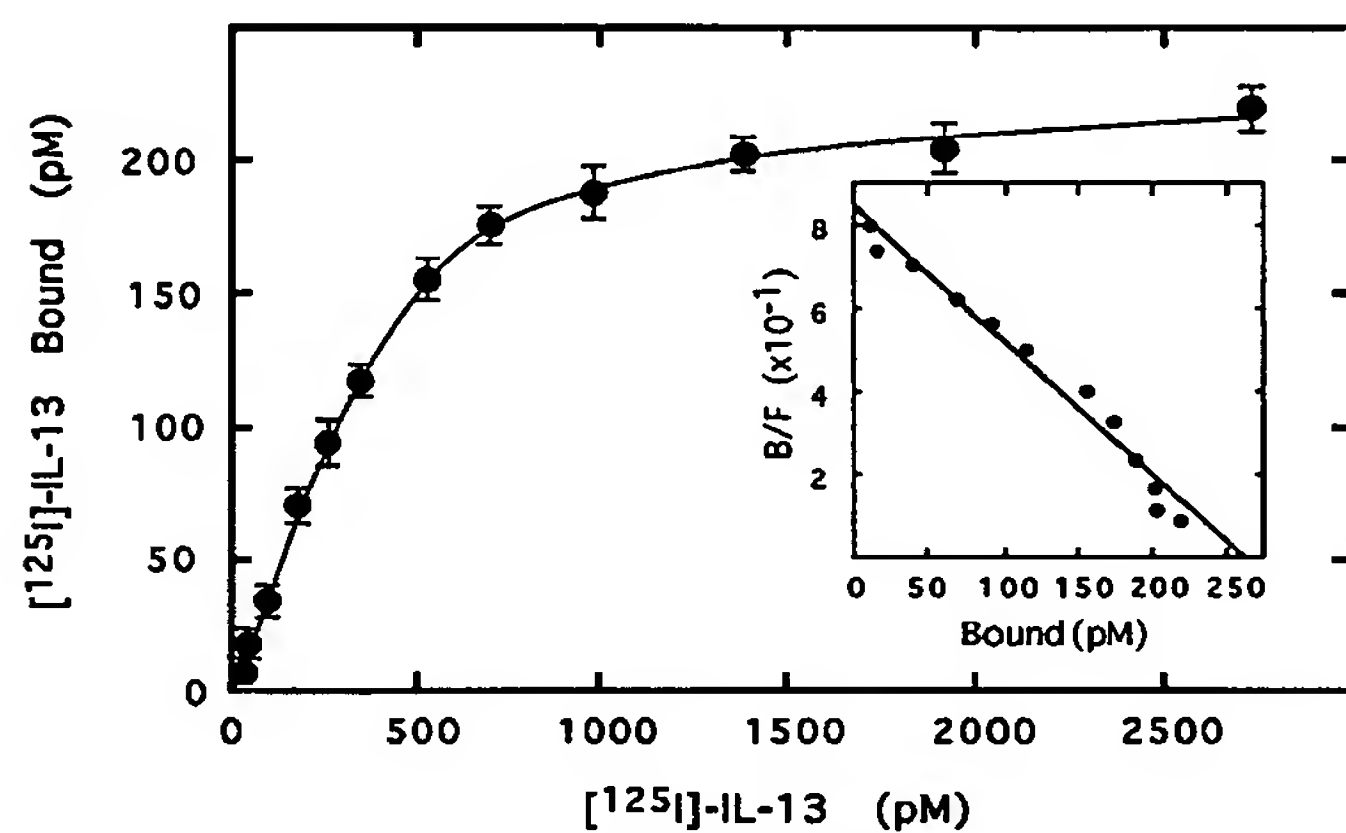


FIG. 5. Characterization of the recombinant IL-13 receptor. COS-7 cells were transfected with the IL-13R cDNA and assayed for binding of  $^{125}\text{I}$ -IL-13. Bars, S.D. ( $n = 3$ ). Inset, Scatchard analysis of the saturation curve, which indicated the presence of  $\sim 560,000$  sites/cell;  $K_d \approx 250$  pM.

lines expressing low number of IL-13 binding sites. Reverse transcription-PCR analysis (Fig. 4) showed that the transcript found in Caki-1 cells is also present, at lower levels, in the keratinocytic A431 cells, premyeloid TF-1 cells, premonocytic U937 cells, and the B-cell line IM9. No transcript was detected in the T-cell line Jurkat. These results are in agreement with our previously reported binding studies (6) and with the known biological targets for IL-13.

**Expression and Characterization of the IL-13 Binding Protein**—COS-7 cells transfected with the isolated cDNA encoding the IL-13R showed specific binding of labeled IL-13. Scatchard analysis of the saturation curve showed a single component site with a  $K_d$  value of  $250 \pm 30$  pM and a maximal binding capacity of  $5.6 \times 10^5$  receptors/cell (Fig. 5). The affinity displayed by the recombinant receptor is in good agreement with the  $K_d$  value of 446 pM for the Caki-1 IL-13 receptor and for that described in several other cells (6).

In competition studies, IL-13 was effective in inhibiting the labeled IL-13 binding to the cloned receptor, with an inhibitory affinity constant ( $K_i$ ) of  $1.5 \pm 0.5$  nM, whereas IL-4 did not inhibit binding (not shown). Cross-linking experiments with labeled IL-13 showed a radioactive band of  $\sim 70$  kDa similar to the one obtained with the Caki-1 cells, shown in Fig. 1b, inset. The radioactivity was completely displaced by a 100-fold excess of IL-13 but not by a 100-fold excess of IL-4 (results not shown), in agreement with the binding experiments. Thus, the pharmacology and the electrophoretic mobility of the covalently cross-linked labeled receptor are those expected from the characterization of the IL-13R present in Caki-1 cells.

TABLE I

$^{125}\text{I}$ -IL-13 binding to COS cells expressing IL-13R, IL-4R, and  $\gamma\text{c}$  alone and in various combinations

COS cells were transfected with the indicated cDNAs and analyzed for  $^{125}\text{I}$ -IL-13 (320 pM) binding 48 h later, as indicated under "Materials and Methods," in the absence or presence of either cold IL-13 (117 nM) or IL-4 (67 nM).

cDNAs transfected	$^{125}\text{I}$ -IL-13, total bound dpm <sup>a</sup>	Bound displaced by	
		IL-13	IL-4
Mock	344	117	116
IL-13R	22,895	22,612	55
IL-4R	427	146	131
IL-13R + IL-4R	21,119	20,942	1772
IL-13R + IL-4R + $\gamma\text{c}$	7769	7481	389
IL-13R + $\gamma\text{c}$	8402	8130	484
IL-4R + $\gamma\text{c}$	351	107	109

<sup>a</sup> Each value is the average of triplicates, the S.D. values of which were less than 10%. The results shown are representative of three independent experiments.

**Co-expression of IL-13R and IL-4R in COS Cells**—We investigated whether the IL-13R and IL-4R would interact in the cell membrane to reconstitute a receptor that allowed cross-competition of both cytokines. The results of the co-expression experiments are shown in Tables I and II. Table I shows the binding of radiolabeled IL-13 to COS cells transfected with cDNAs coding for IL-13R, IL-4R, and  $\gamma\text{c}$  alone and in various combinations. Mock transfected cells displayed low IL-13 binding that is displaced by both IL-13 and IL-4, as described previously (6). Cells transfected with the cDNA for IL-13R showed high IL-13 binding that was completely displaced by IL-13 but not by IL-4; and the expression of IL-4R resulted in IL-13 binding values similar to those observed in mock transfected cells. Thus, IL-13R recognizes IL-13 but not IL-4, and IL-4R does not bind IL-13, as described previously (5, 6). COS cells expressing both IL-13R and IL-4R showed similar binding for IL-13 to that of the cells expressing IL-13R alone; however, about 8–10% of that binding can be displaced by IL-4, suggesting the reconstitution of a binding site shared by both IL-13 and IL-4. The co-expression of  $\gamma\text{c}$  did not improve the reconstitution of shared sites, but resulted in a diminution of IL-13 binding. Similarly, the expression of the  $\gamma\text{c}$  with the IL-13R also resulted in a diminution of IL-13 binding. Work is in progress to investigate an eventual interaction of  $\gamma\text{c}$  and IL-13R that may explain these observations. The co-expression of IL-4R and  $\gamma\text{c}$  did not result in binding sites recognized by IL-13, as described previously (6).

Cells from the same transfection experiments used for binding of labeled IL-13 were also analyzed for binding of labeled IL-4. The binding results with labeled IL-4 (Table II) showed that IL-13 cannot displace labeled IL-4 from the IL-4R alone or from the IL-4R- $\gamma\text{c}$  complex. But when the IL-4R and the IL-13R were co-expressed, the IL-13 was able to displace a fraction of the labeled IL-4 binding, suggesting again an interaction of both chains to reconstitute shared binding sites for both cytokines.

## DISCUSSION

Compelling evidence supports the idea that IL-13 and IL-4 share receptor components as well as signal transduction elements (5, 6, 9–13). However, neither of the two proteins described as components of the high affinity IL-4 receptor, the IL-4R and the  $\gamma\text{c}$ , is responsible for the binding cross-competition of IL-4 and IL-13. Thus, another protein(s), part of the IL-4 complex, is clearly necessary for the recognition of IL-13. Here we describe the characterization and cloning of an IL-13R.

We used Caki-1 cells as source of mRNA for the molecular cloning of the IL-13R because our results suggested that these



TABLE II  
<sup>125</sup>I-IL-4 binding to COS cells expressing IL-13R, IL-4R, and  $\gamma$ c alone and in various combinations

COS cells were transfected with the indicated cDNAs and analyzed for <sup>125</sup>I-IL-4 (160 pM) binding 48 h later, as indicated under "Materials and Methods," in the absence or presence of either cold IL-13 (117 nM) or IL-4 (67 nM).

cDNAs transfected	<sup>125</sup> I-IL-4, total bound	Bound displaced by	
		IL-13	IL-4
	dpm <sup>a</sup>	dpm	
Mock	813	436	416
IL-13R	1159	710	830
IL-4R	20,862	562	20,276
IL-13R + IL-4R	22,008	1793	21,533
IL-13R + IL-4R + $\gamma$ c	27,000	629	26,571
IL-13R + $\gamma$ c	1121	455	445
IL-4R + $\gamma$ c	23,641	88	23,112

<sup>a</sup> Each value is the average of triplicates the S.D. values of which were less than 10%. The results shown are representative of three independent experiments.

cells express an IL-4-IL-13 receptor complex responsible for the binding cross-competition and biological activities of IL-4 and IL-13 similar to that previously described in other cells and, in addition, that these cells overexpress the IL-13R component that binds IL-13.

The cloned cDNA for the IL-13R codes for a protein of 380 amino acids with characteristics of the hematopoietic cytokine receptor family; and the pattern of IL-13R mRNA expression is consistent with the receptor distribution estimated by IL-13 binding and with the known IL-13 target cells.

The new receptor has a short cytoplasmic sequence, which is in line with the observation that probably only the receptor complex that is shared by IL-4 and IL-13 transduces a signal to the cells. However, further work is needed to assess whether this cytoplasmic domain plays a role in associating with other chains and contributes to signal transduction as described for the IL-5R $\alpha$  (22).

The homology of the IL-13R with the IL-5R $\alpha$  chain is interesting because the latter binds IL-5 but needs another protein, the  $\beta$  chain shared with the IL-3 and GM-CSF (granulocyte macrophage colony-stimulating factor) receptors, to form a high affinity receptor complex that is capable of signal transduction (23). However, it should be noted that the IL-13R, when expressed in COS-7 cells, does not need a second chain to reconstitute a high affinity binding site. It therefore resembles the recently described IL-15 binding protein, which also displays high affinity binding for IL-15 in the absence of the two other components of the IL-15 receptor complex (24).

The recombinant receptor recognized IL-13 with high affinity, but was not able to bind IL-4, as expected from the results obtained in Caki-1 cells. The radiolabeled complex observed in the cross-linking experiments with labeled IL-13 has the same mobility as that observed in other cell lines (6). It most probably corresponds to the 60–70-kDa band observed in cross-linking experiments performed with labeled IL-4 (6, 15) in addition to the 140-kDa IL-4R complex, arguing in favor of the interaction of both proteins in the functional receptor complex.

The results of the co-expression of the IL-13R and the IL-4R in COS-7 cells suggest that these proteins can interact with each other in the cell membrane to reconstitute a receptor where IL-13 and IL-4 can compete with each other. However, it should be noted that only a fraction of the expressed IL-4R and IL-13R resulted in shared sites for both cytokines. Several reasons may explain the low yield of reconstituted shared sites. For example, the existence of another protein(s), present in limiting amounts in COS cells, necessary to reconstitute the IL-4-IL-13 receptor complex. Another explanation could be an incorrect stoichiometry of IL-4R and IL-13R in the cell mem-

brane, but this is unlikely because co-transfections producing different relative amounts of the IL-4R and IL-13R did not show major differences in the number of reconstituted receptors. Finally, our results do not exclude the existence of another IL-13R with a better capacity to interact with the IL-4R for the reconstitution of shared binding sites for IL-4 and IL-13. If this is the case, the IL-13R described here could either play a regulatory role in the IL-4-IL-13 receptor complex, or require a protein, different from the known IL-4R, to reconstitute a functionally independent IL-13 receptor complex.

The results obtained in the co-transfection experiments with the  $\gamma$ c demonstrate, as previously suggested (5, 25), that this protein is not the limiting factor for the reconstitution of an IL-4R-IL-13R complex, a conclusion also supported by the absence of detectable  $\gamma$ c mRNA in Caki-1 cells (7). It should be noted that the expression of the  $\gamma$ c improved IL-4 binding as described previously (16), but decreased the binding of IL-13, suggesting complex interaction between the different receptor components. Work is in progress to address this problem.

The cloning of the IL-13R described here should help to address the question of the regulation of the expression of the receptor under normal and pathological conditions, such as allergy, where IL-13 may play a key role. Furthermore, the availability of the cDNA will facilitate the cloning of other proteins necessary for the reconstitution of the IL-4-IL-13 receptor complex or of an eventual independent IL-13 receptor complex and may help the rational design of new drugs that specifically antagonize IL-4 or IL-13 activities.

**Acknowledgments**—We thank Drs. D. Shire and A. Minty for review of the manuscript, Dr. J. Y. Blay for providing the renal carcinoma cells, Dr. W. Sebald for the anti-IL-4R antibodies and the IL-4 antagonist hIL-4.Y124D, and R. Reeb for help with the cell culture.

## REFERENCES

- McKenzie, A. N., Culpepper, J. A., de Waal Malefyt, R., Briere, F., Punnonen, J., Aversa, G., Sato, A., Dang, W., Cocks, B. G., Menon, S., de Vries, J. E., Banchereau, J., and Zurawski, G. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 3735–3739.
- Minty, A., Chalon, P., Derocq, J. M., Dumont, X., Guillemot, J. C., Kaghad, M., Labit, C., Lepatois, P., Liauzun, P., Miloux, B., Minty, C., Casellas, P., Loison, G., Lupker, J., Shire, D., Ferrara, P., and Caput, D. (1993) *Nature* **362**, 248–250.
- Zurawski, G., and de Vries, J. E. (1994) *Immunol. Today* **15**, 19–26.
- Punnonen, J., Aversa, G., Cocks, B. G., McKenzie, A. N., Menon, S., Zurawski, G., de Waal Malefyt, R., and de Vries, J. E. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 3730–3734.
- Zurawski, S. M., Vega, F., Jr., Huyghe, B., and Zurawski, G. (1993) *EMBO J.* **12**, 2663–2670.
- Vita, N., Lefort, S., Laurent, P., Caput, D., and Ferrara, P. (1995) *J. Biol. Chem.* **270**, 3512–3517.
- Obiri, N., Debinski, W., Leonard, W. J., and Puri, R. K. (1995) *J. Biol. Chem.* **270**, 8797–8804.
- Kruse, N., Tony, H. P., and Sebald, W. (1992) *EMBO J.* **11**, 3237–3244.
- Aversa, G., Punnonen, J., Cocks, B. G., de Waal Malefyt, R., Vega, F., Zurawski, S. M., Zurawski, G., and de Vries, J. E. (1993) *J. Exp. Med.* **178**, 2213–2218.
- Lefort, S., Vita, N., Reeb, R., Caput, D., and Ferrara, P. (1995) *FEBS Lett.* **366**, 122–126.
- Zurawski, S. M., Chomarat, P., Djossou, O., Bidaud, C., McKenzie, A. N., Miossec, P., Banchereau, J., and Zurawski, G. (1995) *J. Biol. Chem.* **270**, 13869–13878.
- Keegan, A. D., Johnston, J. A., Tortolani, P. J., McReynolds, L. J., Kinzer, C., O'Shea, J. J., and Paul, W. E. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 7681–7685.
- Wang, L. M., Michieli, P., Lie, W. R., Liu, F., Lee, C. C., Minty, A., Sun, X. J., Levine, A., White, M. F., and Pierce, J. H. (1995) *Blood* **86**, 4218–4227.
- Idzerda, R. L., March, C. J., Mosley, B., Lyman, S. D., Vanden Bos, T., Gimpel, S. D., Din, W. S., Grabstein, K., Widmer, M. B., Park, L. S., Cosman, D., and Backmann, M. P. (1990) *J. Exp. Med.* **171**, 861–873.
- Galizzi, J. P., Castel, B., Djossou, O., Harada, N., Cabrilat, H., Yahia, S. A., Barrett, R., Howard, M., and Banchereau, J. (1990) *J. Biol. Chem.* **265**, 439–444.
- Kondo, M., Takeshita, T., Ishii, N., Nakamura, M., Watanabe, S., Arai, K., and Sugamura, K. (1993) *Science* **262**, 1874–1883.
- Russell, S. M., Keegan, A. D., Harada, N., Nakamura, Y., Noguchi, M., Leland, P., Friedman, M. C., Miyajima, A., Puri, R. K., Paul, W. E., and Leonard, W. J. (1993) *Science* **262**, 1880–1883.
- Caput, D., Beutler, B., Hartog, K., Thayer, R., Brown-Schimer, S., and Cerami, A. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 1670–1674.
- Minty, A., Chalon, P., Guillemot, J.-C., Kaghad, M., Liauzun, P., Magazin, M.,

- Miloux, B., Minty, C., Ramond, P., Vita, N., Lupker, J., Shire, D., Ferrara, P., and Caput, D. (1993) *Eur. Cytokine Netw.* **4**, 99-110
20. Vita, N., Laurent, P., Lefort, S., Chalon, P., Dumont, X., Kaghad, M., Gully, D., Le Fur, G., Caput, D., and Ferrara, P. (1993) *FEBS Lett.* **317**, 139-142
21. Kishimoto, T., Taga, T., and Akira, S. (1994) *Cell* **76**, 253-262
22. Cornelis, S., Fache, I., Van der Heyden, J., Guisez, Y., Tavernier, J., Devos, R., Fiers, W., and Plaetinck, G. (1995) *Eur. J. Immunol.* **25**, 1857-1864
23. Honjo, T. (1991) *Curr. Opin. Cell Biol.* **1**, 201-203
24. Giri, J. G., Kumaki, S., Ahdieh, M., Friend, D. J., Loomis, A., Shanebeck, K., DuBose, R., Cosman, D., Park, L. S., and Anderson, D. M. (1993) *EMBO J.* **14**, 3654-3663
25. Matthews, D. J., Clark, P. A., Herbert, J., Morgan, G., Armitage, R. J., Kinnon, C., Minty, A., Grabstein, K. H., Caput, D., Ferrara, P., and Callard, R. (1995) *Blood* **85**, 38-42

## P-selectin induces the expression of tissue factor on monocytes

ALESSANDRO CELI<sup>\*†</sup>, GIULIANA PELLEGRINI<sup>‡</sup>, ROBERTO LORENZET<sup>‡</sup>, ANTONIO DE BLASI<sup>§</sup>, NEAL READY<sup>\*</sup>, BARBARA C. FURIE<sup>\*</sup>, AND BRUCE FURIE<sup>\*</sup>

<sup>\*</sup>Center for Hemostasis and Thrombosis Research, Division of Hematology–Oncology, New England Medical Center, and Departments of Medicine and Biochemistry, Tufts University School of Medicine, Boston, MA 02111; and <sup>‡</sup>Antonio Taticchi Unit for Atherosclerosis and Thrombosis, and <sup>§</sup>Laboratory of Molecular Biology and Pharmacology of Receptors, Istituto di Ricerche Farmacologiche Mario Negri, Consorzio Mario Negri Sud, 66030 S. Maria Imbaro, Italy

Communicated by Oscar D. Ratnoff, May 23, 1994

**ABSTRACT** P-selectin on activated platelets and stimulated endothelial cells mediates cell adhesion with monocytes and neutrophils. Since activated platelets induce tissue factor on mononuclear leukocytes, we examined the effect of P-selectin on the expression of tissue factor activity in monocytes. Purified P-selectin stimulated tissue factor expression on mononuclear leukocytes in a dose-dependent manner. Chinese hamster ovary (CHO) cells expressing P-selectin stimulated tissue factor procoagulant activity in purified monocytes, whereas untransfected CHO cells and CHO cells expressing E-selectin did not. Anti-P-selectin antibodies inhibited the effects of purified P-selectin and CHO cells expressing P-selectin on monocytes. Incubation of CHO cells expressing P-selectin with monocytes leads to the development of tissue factor mRNA in monocytes and to the expression of tissue factor antigen on the monocyte surface. These results indicate that P-selectin upregulates the expression of tissue factor on monocytes as well as mediates the binding of platelets and endothelial cells with monocytes and neutrophils. The binding of P-selectin to monocytes in the area of vascular injury may be a component of a mechanism that initiates thrombosis.

Blood clotting is a host defense mechanism that, in parallel with the inflammatory and repair responses, preserves the integrity of the vascular system after tissue injury (1). Platelets, leukocytes, and endothelial cells are among the cellular components critical for this process. The plasma blood clotting proteins participate in a molecular cascade in which tissue injury activates blood coagulation, leading to the formation of a fibrin clot (2). The response to vascular injury culminates in the formation of a platelet plug, the deposition of leukocytes in injured tissue, and the initiation of inflammation and wound healing. Blood coagulation is initiated through the action of tissue factor. Normally not exposed to blood, tissue factor is an integral membrane protein expressed constitutively on the surface of nonvascular cells. Monocytes and endothelial cells can be induced to express tissue factor on their surface (3, 4). Monocytes can be activated by endotoxin (5), immune complexes (6), certain cytokines (7), and platelets (8–11), leading to tissue factor expression, whereas endothelial cells express tissue factor when stimulated by certain cytokines (12). Tissue factor, with a molecular weight of 43,000 (13, 14), binds factor VII and factor VIIa to form the tissue factor/factor VIIa complex that activates factor IX and factor X (15).

P-selectin is a cell adhesion molecule that mediates the interaction of platelets and endothelial cells with neutrophils and monocytes (16, 17). P-selectin, a member of the selectin family of adhesion molecules (18–21), is an integral membrane protein found in the  $\alpha$  granules of platelets (22, 23) and the Weibel–Palade bodies of endothelial cells (24, 25). Upon

stimulation of these cells by agonists such as thrombin, P-selectin is phosphorylated (26, 27) and rapidly translocated to the plasma membrane (23). P-selectin is a lectin that binds to lineage-specific carbohydrates on the surface of monocytes and neutrophils (28–30). This protein binds to a mucin-like glycoprotein PSGL-1 that must be properly glycosylated to retain functional properties as the P-selectin ligand (31). P-selectin on platelets mediates the accumulation of leukocytes into the growing thrombus during experimental thrombosis *in vivo* (32). Inhibitory antibodies that block the interaction of P-selectin on platelets with the P-selectin ligand on leukocytes inhibit the uptake of leukocytes into the thrombus and inhibit the magnitude of thrombus formation. These experiments have demonstrated that P-selectin mediates monocyte and neutrophil interaction with activated platelets *in vitro* and *in vivo*. The potential exists for stimulation of leukocyte effector function by P-selectin binding. To evaluate this potential, we have examined the ability of P-selectin to upregulate tissue factor expression on monocytes. In the current study, we demonstrate that P-selectin induces the expression of tissue factor on monocytes exposed to P-selectin.

### EXPERIMENTAL PROCEDURES

**Proteins and Cell Lines.** P-selectin was purified from platelets by detergent extraction of platelet membranes, heparin-Sepharose chromatography, and affinity chromatography using AC1.2 (30). P-selectin appeared homogeneous by SDS gel electrophoresis. Production of the noninhibitory anti-P-selectin antibody AC1.2 and the inhibitory P-selectin antibody GA6 have been described (16, 32). The inhibitory anti-tissue factor antibody HTF1 (33) was the gift of Yale Nemerson (Mt. Sinai School of Medicine). Chinese hamster ovary (CHO) cells expressing P-selectin (CHO:P-selectin) or E-selectin (CHO:E-selectin) adhere to leukocytes, in contrast to CHO cells (30).

**Cell Isolation and Culture.** Blood was obtained from normal volunteers and anticoagulated with 0.1 vol of 3.8% sodium citrate/0.15 M NaCl. Erythrocytes and leukocytes were sedimented at  $150 \times g$  for 15 min at 10°C. Platelet-rich plasma was used for platelet isolation (11). The sedimented erythrocytes and leukocytes were adjusted to the original volume with 0.38% sodium citrate/0.15 M NaCl, sedimented, and layered onto Ficoll/Hypaque (Pharmacia). After centrifugation at  $400 \times g$  for 30 min at 8°C, mononuclear cells were removed, diluted in 0.38% sodium citrate/0.15 M NaCl, and sedimented at  $450 \times g$  for 7 min at 8°C. The pellet was resuspended in buffer and the cells were washed three times. Mononuclear cells were resuspended in RPMI 1640 medium

Abbreviations: GAPDH, glyceraldehyde phosphate dehydrogenase; LPS, lipopolysaccharide.

<sup>†</sup>Present address: Consiglio Nazionale delle Ricerche Institute of Clinical Physiology and 2a Clinica Medica, University of Pisa, 56126 Pisa, Italy.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.



supplemented with penicillin (100 units/ml) and streptomycin (100  $\mu$ g/ml) at a cell concentration of  $10^7$  cells per ml. Cell viability (>94%) was assessed by trypan blue exclusion. Platelet contamination was 0.3–0.5 platelet per mononuclear cell. The monocytes in this population, as assessed by nonspecific esterase staining, were 25–30%. Monocytes were purified by using a discontinuous Percoll density gradient. Mononuclear cells resuspended in RPMI 1640 medium supplemented with 10% heat-inactivated fetal calf serum were layered onto a 46% isosmotic Percoll solution. After centrifugation at 4°C for 30 min at  $550 \times g$ , the cells at the RPMI 1640 medium/Percoll interface were collected and washed twice at  $440 \times g$  in serum-free RPMI 1640 medium. These preparations contained about 85% monocytes and 15% lymphocytes. Mononuclear cells were incubated with platelets, purified P-selectin, or CHO cells at 37°C in sterile, pyrogen-free stoppered test tubes or in tissue culture plates under 5% CO<sub>2</sub>/95% air.

**Control of Endotoxin Contamination.** All reagents used for cell isolation and culture were prepared with endotoxin-free water. Solutions were prepared in glassware rendered endotoxin-free by high temperature. All reagents showed a level of endotoxin contamination, as assessed by a chromogenic *Limulus* assay (BioWhittaker), lower than 0.1 endotoxin unit/ml. All experiments were performed under sterile conditions.

**Coagulant Activity Assay.** Except where otherwise indicated, cells were disrupted by three freeze–thaw cycles prior to measurement of procoagulant activity by a one-stage clotting assay (11). In some experiments, intact cells were used. Intact or disrupted cells (100  $\mu$ l) were mixed with 100  $\mu$ l of normal human plasma at 37°C. After 30 sec, 100  $\mu$ l of 25 mM CaCl<sub>2</sub> at 37°C was added to the mixture and the time to clot formation was recorded. The values were converted to arbitrary units of procoagulant activity by comparison with a standard curve obtained by using a human brain thromboplastin standard (U.K. 1) from the reference laboratory (gift of L. Poller, Manchester, U.K.). This preparation was assigned a value of 1000 units for a clotting time of 20 sec. All experimental results shown are the averages of duplicate or triplicate independent cultures within one experiment.

**Quantitative PCR.** Oligonucleotide primers corresponding to bp 178–198 (sense) and bp 495–515 (antisense) of the human tissue factor coding sequence and bp 64–86 (sense) and bp 581–603 (antisense) in the coding sequence of human glyceraldehyde phosphate dehydrogenase (GAPDH) were synthesized. Quantitative PCR was performed with 7  $\mu$ l of cDNA from 100  $\mu$ l of total cDNA resulting from reverse transcription of 1  $\mu$ g of RNA. The assay mixture contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.4  $\mu$ g of each primer, 250  $\mu$ M each dNTP, 2.5 units of *Taq* polymerase. The amplification conditions were 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min. After 25 cycles, the PCR products from GAPDH mRNA (528 bp) and tissue factor mRNA (317 bp) were analyzed by electrophoresis in a 2% agarose gel stained with ethidium bromide. For tissue factor, Southern blot analysis was performed on 8  $\mu$ l of the PCR product using the labeled 317-bp fragment as a probe. The relative intensity of the bands visualized by autoradiography was measured by laser densitometry (34).

**Fluorescence Microscopy.** Mononuclear cells were grown on glass slides in the presence of buffer, lipopolysaccharide (LPS) (*Escherichia coli* 055:B5W; Difco), or P-selectin for 6 hr. The cells were fixed with 2% paraformaldehyde for 1 hr at 4°C and the reaction was quenched with TBS/10 mM NH<sub>4</sub>Cl for 10 min. After incubation with 2% bovine serum albumin/phosphate-buffered saline (BSA/PBS) for 30 min and 2% BSA/PBS/0.2% human immunoglobulin for 30 min, the cells were incubated with HTF1 (5  $\mu$ g/ml in 2% BSA/PBS/0.2% human immunoglobulin for 30 min at 23°C. After

washing, fluorescein isothiocyanate-conjugated anti-mouse immunoglobulin was reacted with the cells for 30 min at 23°C. The cells were examined with a Zeiss fluorescence microscope.

## RESULTS

The incubation of platelets with mononuclear leukocytes leads to expression of procoagulant activity. Mononuclear cells or platelets alone incubated for 16 hr at 37°C expressed minimal procoagulant activity. However, when platelets and mononuclear cells were coincubated under these conditions, we observed a 50- to 100-fold increase in procoagulant activity, confirming earlier results (8). This procoagulant activity is due to tissue factor inasmuch as stimulated mononuclear cells incubated with the inhibitory anti-tissue factor antibody HTF1 (30  $\mu$ g/ml) expressed no procoagulant activity.

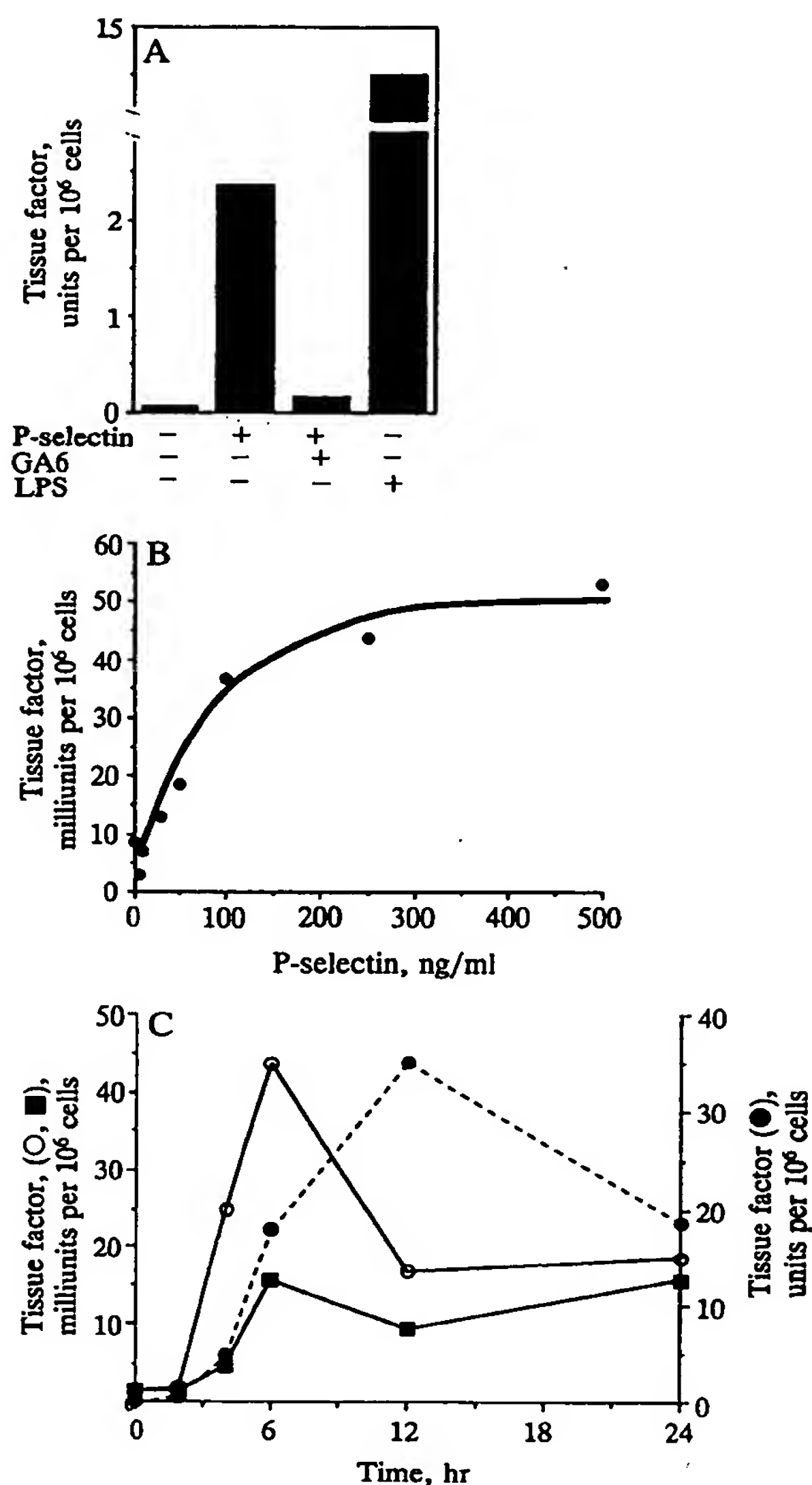
Purified P-selectin induced tissue factor expression by 50- to 100-fold under the conditions used (Fig. 1A). Tissue factor expression was inhibited if GA6, an inhibitory monoclonal P-selectin antibody, was added to cells prior to P-selectin exposure. This P-selectin preparation did not have detectable endotoxin, as monitored in the *Limulus* assay. However, endotoxin is a well-known potent stimulator of tissue factor expression in monocytes (5, 35). To distinguish between upregulation of tissue factor due to P-selectin and that due to contaminating endotoxin that may be present below the levels of detection of the *Limulus* assay, P-selectin was heated at 100°C for 30 min to denature the protein. Heat denaturation of P-selectin abolished the upregulation of tissue factor activity. In contrast, LPS, in its native form and after heating at 100°C for 30 min, stimulated tissue factor activity. These results indicate that P-selectin, and not minor endotoxin contaminants, upregulates tissue factor activity on monocytes. Intact monocytes stimulated by either P-selectin or endotoxin expressed  $\approx 15\%$  of the tissue factor activity of lysed cells. This is similar to the results of Drake *et al.* (36) and Levy *et al.* (37), where the tissue factor activity of intact cells was 18–21% that of lysed cells. These results indicate that the tissue factor antigen expressed on the cell surface after P-selectin exposure (see below) is functional.

The expression of tissue factor activity as a function of P-selectin concentration was studied. Expression of tissue factor activity is dependent on P-selectin concentration; the response was saturable (Fig. 1B). This activity was inhibited with an inhibitory anti-tissue factor antibody.

Unstimulated monocytes do not contain tissue factor in a storage pool nor is any tissue factor expressed on the cell surface (36). After stimulation, the kinetics of tissue factor expression in monocytes induced by P-selectin was compared to that induced by LPS. Tissue factor expression induced by P-selectin was observed at 4 hr, peaked at 6 hr, and then gradually decayed (Fig. 1C). The peak of P-selectin-induced tissue factor preceded the LPS-induced peak of tissue factor. The low level of tissue factor activity measured in the absence of P-selectin or LPS represents a low level of spontaneous activation of monocytes in culture. Although the endotoxin level in the tissue culture medium is below the sensitivity limit of the amebocyte lysate assay, we cannot eliminate a small amount of endotoxin as the cause of monocyte activation.

Cells expressing P-selectin also mediate the induction of tissue factor expression; we compared CHO:P-selectin with untransfected CHO cells for their ability to induce tissue factor expression on monocytes. CHO cells incubated with mononuclear cells had no effect on the development of tissue factor activity (Fig. 2A). However, CHO:P-selectin cells incubated with mononuclear cells yielded approximately a 10-fold increase in tissue factor activity. This activation was

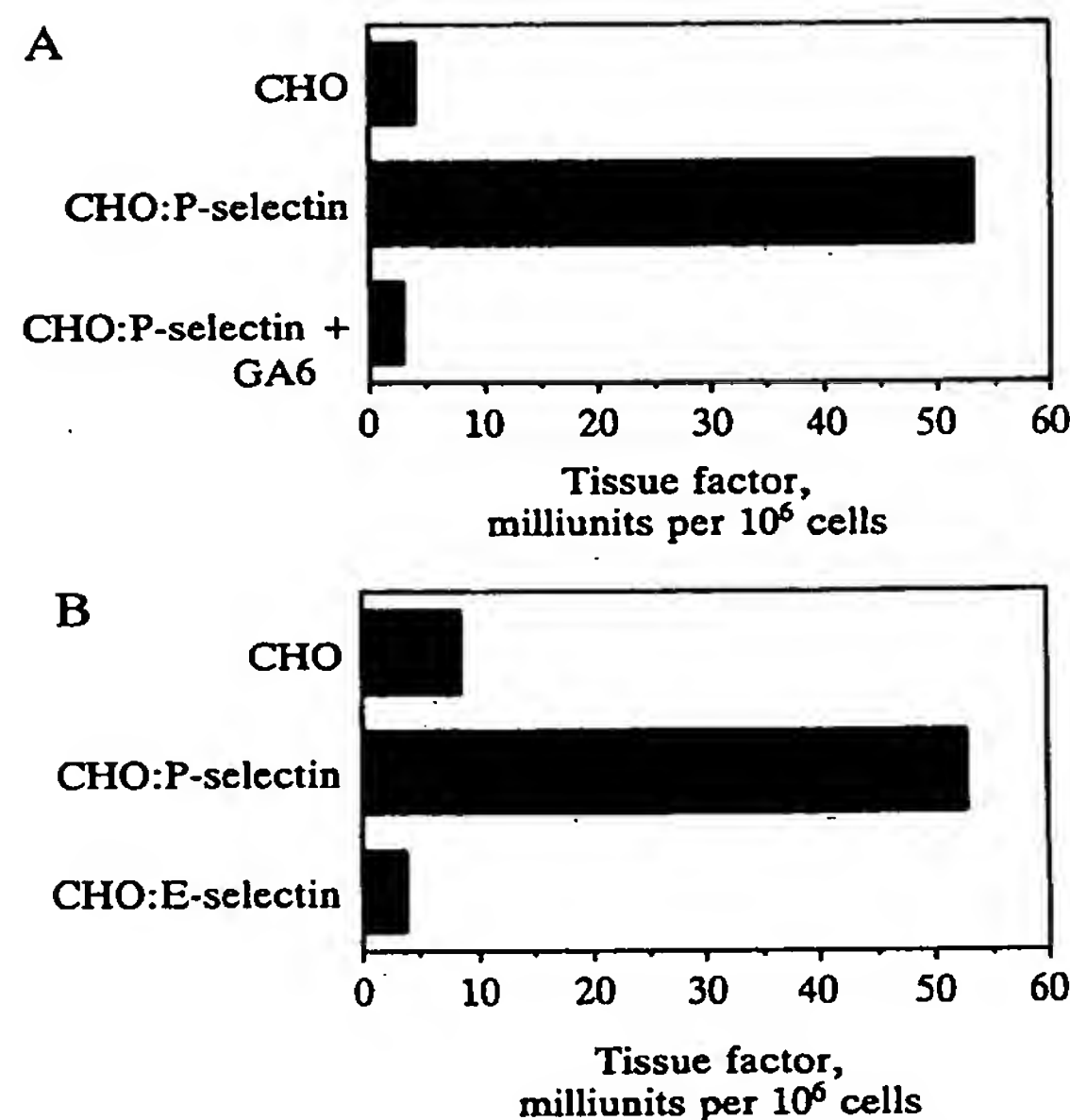




**FIG. 1.** Effect of purified P-selectin on tissue factor expression in monocytes. (A) Tissue factor expression was measured in monocytes that were untreated or exposed to P-selectin (250 ng/ml), P-selectin plus GA6 (100  $\mu$ g/ml), or LPS (100 ng/ml) for 6 hr at 37°C. (B) Effect of increasing concentrations of purified P-selectin on tissue factor expression. Mononuclear cells were incubated with P-selectin at various concentrations in stoppered, endotoxin-free tubes at 37°C. After 6 hr, cells were assayed for tissue factor activity. (C) Kinetics describing the development of tissue factor expression in mononuclear cells as induced by P-selectin. Mononuclear cells were incubated with buffer, P-selectin (250 ng/ml), or LPS (100 ng/ml). After the indicated time, the cells were assayed for tissue factor activity at various time intervals. ○, P-selectin; ●, LPS; ■, buffer control.

blocked in the presence of the anti-P-selectin antibody GA6. While both isolated P-selectin and CHO:P-selectin cells induce tissue factor in monocytes, activated platelets elicit significantly higher levels of tissue factor when incubated with monocytes.

Although among mononuclear leukocytes the monocyte is the only cell capable of expressing tissue factor (38), we confirmed that the cell responsible for P-selectin-induced upregulation of tissue factor expression was the monocyte. Monocytes isolated from mononuclear cells with a Percoll density gradient, when incubated with CHO:P-selectin, also



**FIG. 2.** CHO:P-selectin induces expression of tissue factor in monocytes. (A) CHO cells, either naive (CHO) or CHO:P-selectin, were cultured to confluency. Mononuclear cells were added and incubated for 6 hr in the presence or absence of F(ab')<sub>2</sub> fragments of GA6 (100  $\mu$ g/ml) before assay for tissue factor activity. (B) CHO, CHO:P-selectin, or CHO:E-selectin cells were cultured as described above. Percoll gradient-purified monocytes were incubated as described above before assay for tissue factor activity.

expressed tissue factor activity (Fig. 2B). We asked whether the monocyte stimulation induced by CHO:P-selectin was due to P-selectin or whether the binding of monocytes to CHO cells mediated by any cell adhesion molecule might elicit adhesion-dependent cell activation. Monocytes were incubated with CHO:E-selectin cells (30). Although monocytes express an E-selectin ligand, which is distinct from the P-selectin ligand (30, 31, 39), monocytes bound to CHO:E-selectin cells did not express tissue factor (Fig. 2B).

To determine the steady-state levels of tissue factor mRNA, we performed quantitative reverse transcriptase PCR (40, 41). RNA prepared from untreated control cells, cells treated with LPS, P-selectin, or P-selectin plus the GA6 antibody was reverse transcribed and used for parallel assay of tissue factor mRNA and GAPDH PCR amplification. The expected PCR product for tissue factor (317 bp) was obtained and its identity was confirmed by restriction analysis with *Acc I* and *Rsa I*. To quantify the effect of P-selectin on tissue factor mRNA levels, we performed a Southern blot analysis of the PCR products during amplification. The PCR product from the control cells is barely visible, indicating that the expression of tissue factor mRNA is negligible (Fig. 3). The level of tissue factor mRNA was increased by ~50-fold in cells exposed to LPS or to P-selectin. GA6 suppressed the level of tissue factor mRNA. Control experiments measuring the levels of GAPDH mRNA indicated similar mRNA levels in control and P-selectin- and LPS-treated cells, indicating that the efficiency of reverse transcription was comparable among the experimental groups. For both tissue factor and GAPDH, an identical PCR performed in parallel without including reverse transcriptase gave no amplification products, thus ruling out PCR carryover.

To confirm that P-selectin upregulates tissue factor expression on the surface of monocytes, mononuclear cells were cultured on glass slides and incubated for 6 hr with P-selectin, LPS, or buffer. The adherent monocytes incu-

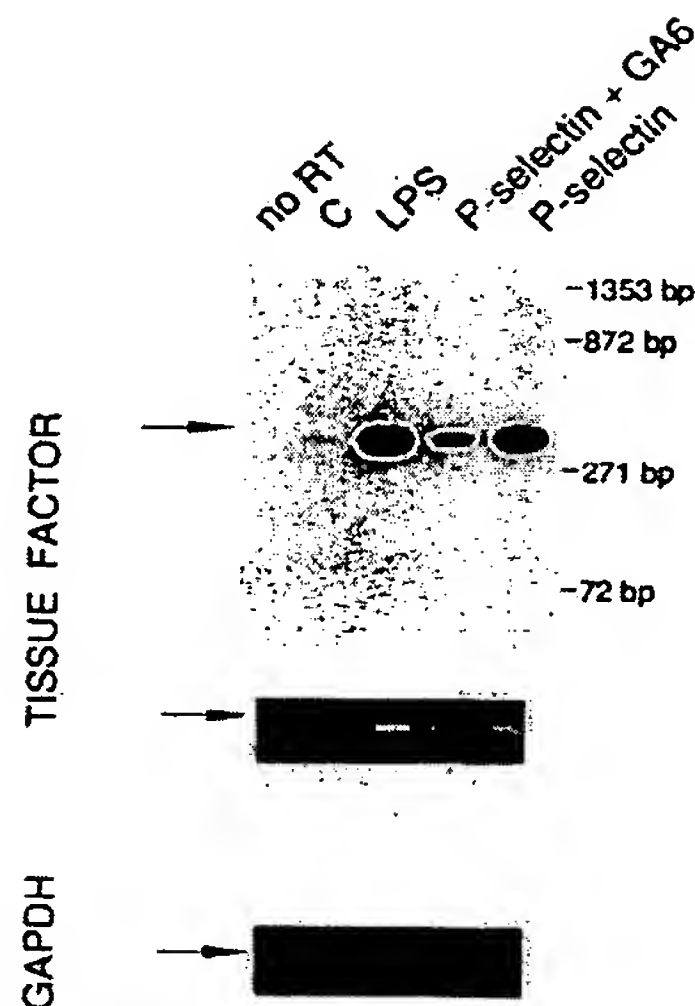


FIG. 3. Effect of P-selectin stimulation on tissue factor mRNA levels. Reverse transcriptase PCR analysis of tissue factor and GAPDH mRNA expression in monocytes untreated (C) or exposed to LPS (100 ng/ml), P-selectin (250 ng/ml) plus GA6 (100  $\mu$ g/ml), or P-selectin alone for 6 hr at 37°C. (Top) Analysis of tissue factor mRNA by Southern blotting. Tissue factor cDNA (317 bp) is indicated by the arrow and the position of the molecular weight markers is indicated. (Middle) Ethidium bromide staining of the same PCR products. The specific tissue factor cDNA product (317 bp) is indicated by the arrow. (Bottom) Ethidium bromide staining of reaction products. The GAPDH cDNA product (528 bp) is indicated by the arrow. A parallel reaction without reverse transcriptase was carried out to test possible contamination (no RT).

bated with P-selectin or LPS were reactive with an anti-tissue factor antibody, while control monocytes incubated with buffer were not (Fig. 4).

### DISCUSSION

The initiation of blood coagulation following tissue injury requires the exposure of tissue factor on the cell surface to plasma factor VIIa, a vitamin K-dependent clotting enzyme that circulates at low concentration in the blood (42). The tissue factor/factor VIIa complex on the cell surface activates both factor IX and factor X, thus leading to the sequential conversion of the blood clotting proenzymes to enzymes and the generation of a fibrin clot. Tissue factor plays a pivotal role in this process insofar as the expression of tissue factor activity is required for initiation and propagation of the signals necessary for clot formation. Nonvascular cells express tissue factor constitutively. During tissue injury, these cells make contact with flowing blood, thus initiating blood clotting on a rapid time scale. In contrast, monocytes and endothelial cells do not express tissue factor constitutively but only after cell stimulation (12, 43, 44). *De novo* synthesis of tissue factor by cells in culture requires at least 2 hr before tissue factor activity is expressed on the cell surface (43).

P-selectin is an adhesion molecule on activated platelets and stimulated endothelial cells that mediates the binding of certain leukocytes to these cells. The P-selectin ligand expressed on these cells include the Le<sup>x</sup> carbohydrate structure, sialic acid, and a protein component PSGL-1 (31). The binding of P-selectin to leukocytes may also up- or down-regulate cell functions. Since monocytes express the P-selectin ligand and synthesize tissue factor after cell stimulation, we have evaluated whether P-selectin binding can lead to the expression of tissue factor on monocytes.

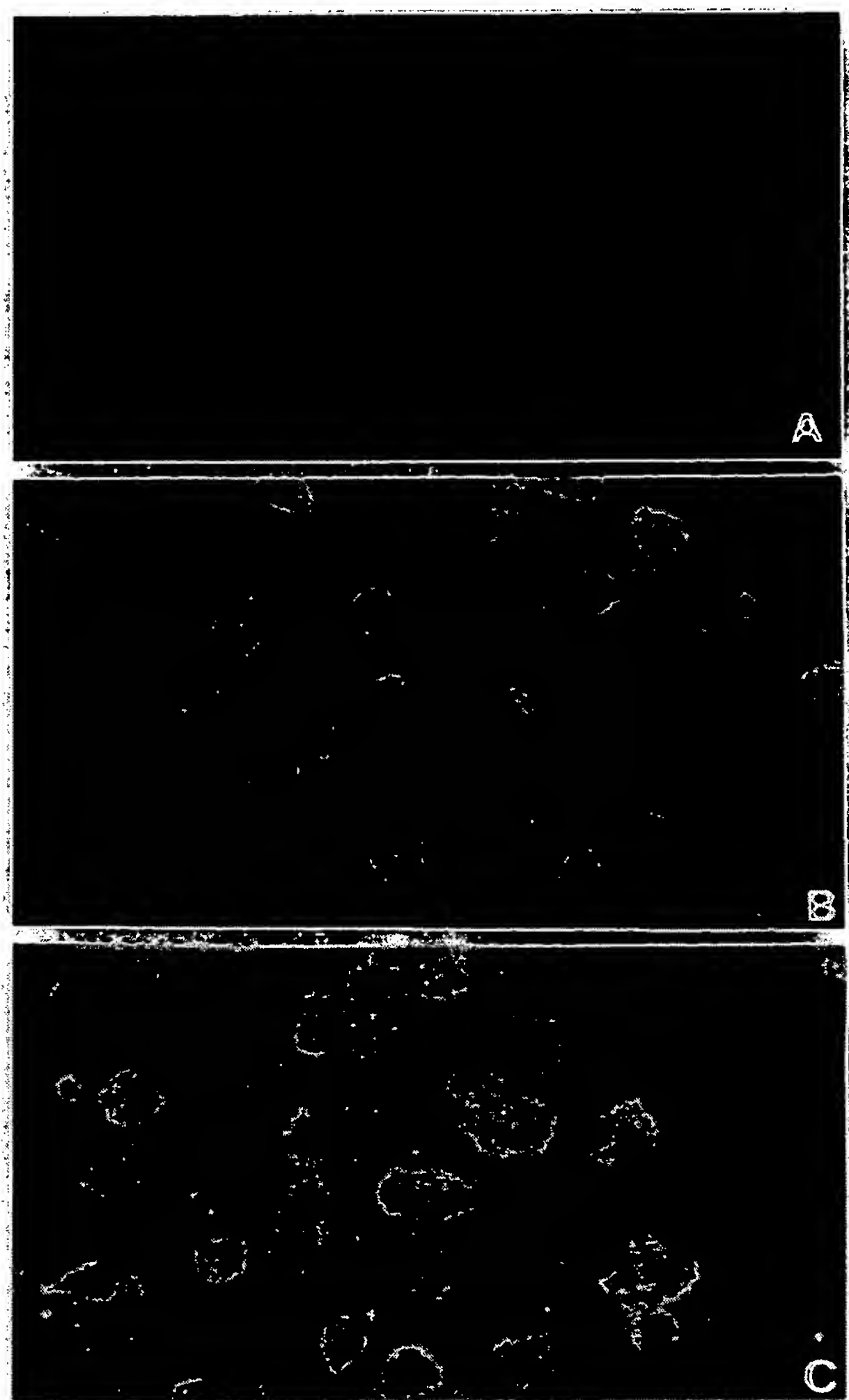


FIG. 4. Immunofluorescence staining of P-selectin-stimulated monocytes with a tissue factor antibody. Mononuclear cells were cultured on glass slides and stimulated with P-selectin, LPS, or buffer control for 6 hr. The cells were stained with fluorescein isothiocyanate-conjugated HTF1, an anti-tissue factor antibody. (A) Buffer control. (B) P-selectin. (C) LPS.

We have demonstrated that P-selectin in three distinct forms is capable of upregulating tissue factor: (i) platelets, which express P-selectin on the cell surface upon activation; (ii) CHO cells stably transfected with the P-selectin; (iii) purified P-selectin. Interaction of platelets or CHO:P-selectin with monocytes through P-selectin appears sufficient to initiate signal transduction for tissue factor expression since inhibitory antibodies to P-selectin block cell-cell interaction and the tissue factor response. However, activated platelets induce significantly more tissue factor expression than do either P-selectin or CHO:P-selectin cells. While P-selectin plays an important role in tethering neutrophils to activated endothelial cells, activation of the neutrophils, as measured by increases in cytosolic  $Ca^{2+}$ , upregulation of CD11/CD18, change in morphology, or priming for granule secretion, requires participation of a second molecule, platelet-activation factor, on the surface of activated endothelial cells (45). It is likely that in addition to P-selectin the activated platelet provides a second ligand either soluble or on the surface of the platelet that augments the synthesis of tissue factor by bound monocytes.



The potency of P-selectin alone in upregulating the synthesis of tissue factor by monocytes is small compared to LPS. However, the relative low potency of P-selectin as measured in our experiments does not exclude a physiologic role for P-selectin in thrombogenesis. For example, we have recently demonstrated in an *ex vivo* model that fibrin deposition after leukocyte accumulation on a vascular graft is inhibited by anti-P-selectin antibodies (32). Low levels of tissue factor expression as mediated by P-selectin may be significant *in vivo* in fibrin generation during vessel wall injury. Alternatively, P-selectin may be just one component required for the initiation of tissue factor expression; for example, a second molecule (e.g., 12-HETE) significantly amplifies the signal for tissue factor expression (unpublished results).

Blood coagulation is a host defense mechanism that involves an immediate, rapid response to the loss of integrity of a closed vascular system and a slower, delayed response that leads to thrombogenesis, inflammation, and wound healing. Tissue factor on nonvascular cells, which becomes accessible to flowing blood during vessel injury, is responsible for the immediate response. In contrast, the delayed expression of tissue factor in the region of tissue injury may be critical to maintenance of thrombus. In our proposed model, neutrophils and monocytes accumulate in the area of tissue injury due to the expression of P-selectin on activated platelets that collect in the subendothelium and on stimulated endothelium in the injured vasculature. Upon interaction with P-selectin, and possibly other required ligands, on platelets and endothelial cells, monocytes upregulate the *de novo* synthesis of tissue factor. Within several hours, tissue factor expression on the cell surface supports and maintains the local activation of blood coagulation and the deposition of fibrin. It is likely that P-selectin initiates a general cell response, of which tissue factor biosynthesis is but one of a series of metabolic and synthetic pathways activated. P-selectin-mediated effector function in leukocytes will likely play an important role in inflammation and thrombosis.

A.C. and G.P. made equivalent contributions to this work. We thank Dr. Susan Sajer for purified P-selectin and Dr. Yale Nemerson for the inhibitory anti-tissue factor antibody HTF1. This work was supported by Grant HL42443 from the National Institutes of Health and by the Italian National Research Council (Consiglio Nazionale delle Ricerche, Rome), Progetto finalizzato FATMA, Contract n.94.00694.41, 1994. N.R. was the recipient of a National Research Service Award (T32 HL07437) from the National Institutes of Health.

1. Furie, B. & Furie, B. C. (1988) *Cell* 53, 505–518.
2. Furie, B. & Furie, B. C. (1992) *N. Engl. J. Med.* 326, 800–806.
3. Niemetz, J. (1972) *J. Clin. Invest.* 51, 307–313.
4. Colucci, M., Balconi, G., Lorenzet, R., Pietra, A., Locati, D., Donati, M. B. & Semeraro, N. (1983) *J. Clin. Invest.* 71, 1893–1896.
5. Semeraro, N., Biondi, A., Lorenzet, R., Locati, D., Mantovani, A. & Donati, M. B. (1983) *Immunology* 50, 529–535.
6. Rothberger, H., Zimmermann, T. S., Spielberg, H. L. & Vaughan, J. H. (1977) *J. Clin. Invest.* 59, 549–557.
7. Conkling, P. R., Greenberg, C. S. & Weinberg, J. B. (1988) *Blood* 72, 128–133.
8. Niemetz, J. & Marcus, A. J. (1974) *J. Clin. Invest.* 54, 1437–1443.
9. Osterud, B. & Bjorklid, E. (1982) *Scand. J. Hematol.* 29, 175–187.
10. Pinder, P. B., Hunt, J. A. & Zacharsky, L. R. (1985) *Am. J. Hematol.* 19, 317–325.
11. Lorenzet, R., Niemetz, J., Marcus, A. J. & Broekman, M. J. (1985) *J. Clin. Invest.* 76, 418–423.
12. Bevilacqua, M. P., Pober, J. S., Majesau, G. R., Cotran, R. S. & Gimbrone, M. A. (1984) *J. Exp. Med.* 160, 618–623.
13. Spicer, E. K., Horton, R., Bloem, L., Bach, R., Williams, K. R., Guha, A., Kraus, J., Lin, T.-C., Nemerson, Y. & Konigsberg, W. H. (1987) *Proc. Natl. Acad. Sci. USA* 84, 5148–5152.
14. Morrissey, J. H., Fakhrai, H. & Edgington, T. S. (1987) *Cell* 50, 129–135.
15. Osterud, B. & Rapaport, S. I. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5260–5264.
16. Larsen, E., Celi, A., Gilbert, G. E., Furie, B. C., Erban, J. K., Bonfanti, R., Wagner, D. D. & Furie, B. (1989) *Cell* 59, 305–312.
17. Geng, J. G., Bevilacqua, M. P., Moore, K. L., McIntyre, T. M., Prescott, S. M., Kim, J. M., Bliss, G. A., Zimmerman, G. A. & McEver, R. P. (1990) *Nature (London)* 343, 757–760.
18. Johnston, G. I., Cook, R. G. & McEver, R. P. (1989) *Cell* 56, 1033–1044.
19. Bevilacqua, M. P., Stengalin, S., Gimbrone, M. A., Jr., & Seed, B. (1989) *Science* 243, 1160–1165.
20. Lasky, L. A., Singer, M. S., Yednock, T. A., Dowbenko, D., Fennie, C., Rodriguez, H., Nguyen, T., Stachel, S. & Rosen, S. D. (1989) *Cell* 56, 1045–1055.
21. Tedder, T. F., Isaacs, C. M., Ernst, T. J., Demetri, G. D., Adler, D. A. & Distech, C. M. (1989) *J. Exp. Med.* 170, 123–133.
22. Stenberg, P. E., McEver, R. P., Shuman, M. A., Jacques, Y. V. & Bainton, D. F. (1985) *J. Cell Biol.* 101, 880–886.
23. Berman, C. L., Yeo, E., Wencel-Drake, J. D., Furie, B. C., Ginsberg, M. H. & Furie, B. (1986) *J. Clin. Invest.* 78, 130–137.
24. Bonfanti, R., Furie, B. C., Furie, B. & Wagner, D. D. (1989) *Blood* 73, 1109–1112.
25. McEver, R. P., Beckstead, J. H., Moore, K. L., Marshall-Carlson, L. & Bainton, D. F. (1989) *J. Clin. Invest.* 84, 92–99.
26. Crovella, C. S., Furie, B. C. & Furie, B. (1993) *J. Biol. Chem.* 268, 14590–14593.
27. Fujimoto, T. & McEver, R. P. (1993) *Blood* 82, 1758–1766.
28. Larsen, E., Palabrica, T., Sajer, S., Gilbert, G. E., Wagner, D. D., Furie, B. C. & Furie, B. (1990) *Cell* 63, 467–474.
29. Polley, M. J., Phillips, M. L., Wayner, E., Nudelman, E., Singhal, A. K., Hakomori, S.-I. & Paulson, J. C. (1991) *Proc. Natl. Acad. Sci. USA* 88, 6224–6228.
30. Larsen, G. R., Sako, D., Ahern, T. J., Shafer, M., Erban, J., Sajer, S. A., Gibson, R. M., Wagner, D. D., Furie, B. C. & Furie, B. (1992) *J. Biol. Chem.* 267, 11104–11110.
31. Sako, D., Chang, X.-J., Barone, K. M., Vachino, G., White, H. M., Shaw, G., Veldman, T., Bean, K. M., Furie, B., Ahern, T. J., Cumming, D. A. & Larsen, G. R. (1993) *Cell* 75, 1179–1186.
32. Palabrica, T., Lobb, R., Furie, B. C., Aronowitz, M., Benjamin, C., Hsu, Y.-M., Sajer, S. A. & Furie, B. (1992) *Nature (London)* 359, 848–851.
33. Carson, S. D., Ross, S. E., Bach, R. & Guha, A. (1987) *Blood* 70, 490–493.
34. Parruti, G., Perracchia, F., Sallese, M., Ambrosini, M., Masini, M. T., Rotilio, D. & De Blasi, A. (1993) *J. Biol. Chem.* 268, 9753–9761.
35. Rickles, F. R., Rick, P. D. & Why, M. V. (1977) *J. Clin. Invest.* 59, 1188–1195.
36. Drake, T. A., Ruf, W., Morrissey, J. H. & Edgington, T. S. (1989) *J. Cell Biol.* 109, 389–395.
37. Levy, G. A., Schwartz, B. S., Curtiss, L. K. & Edgington, T. S. (1985) *J. Clin. Invest.* 76, 548–555.
38. Drake, T. A., Morrissey, J. H. & Edgington, T. S. (1989) *Am. J. Pathol.* 134, 1087–1097.
39. Levinovitz, A., Muhlhoff, J., Isenmann, S. & Vestweber, D. (1993) *J. Cell Biol.* 121, 449–459.
40. Wang, A. M., Doyle, M. V. & Mark, D. F. (1989) *Proc. Natl. Acad. Sci. USA* 86, 9717–9721.
41. Ungerer, M., Bohm, M., Elce, J. S., Erdmann, E. & Lohse, M. J. (1993) *Circulation* 87, 454–463.
42. Morrissey, J. H., Macik, B. G., Neuenschwander, P. F. & Comp, P. C. (1993) *Blood* 81, 734–744.
43. Edwards, R. L. & Rickles, F. R. (1984) *Prog. Hemostasis Thromb.* 7, 183–209.
44. Gregory, S. A. & Edgington, T. S. (1985) *J. Clin. Invest.* 76, 2440–2445.
45. Lorant, D. E., Topham, M. K., Whatley, R. R., McEver, R. P., McIntyre, T. M., Prescott, S. M. & Zimmerman, G. A. (1993) *J. Clin. Invest.* 92, 559–570.

# The P-selectin Glycoprotein Ligand from Human Neutrophils Displays Sialylated, Fucosylated, O-Linked Poly-N-acetyllactosamine\*

(Received for publication, February 1, 1994, and in revised form, June 6, 1994)

Kevin L. Moore†§¶, Steven F. Eaton||, David E. Lyons\*\*, Henri S. Lichenstein\*\*, Richard D. Cummings††, and Rodger P. McEver†§¶§§

From the Departments of †Medicine and ††Biochemistry and the §§W. K. Warren Medical Research Institute, University of Oklahoma Health Sciences Center and the §Cardiovascular Biology Research Program, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma 73104, the ||Department of Biochemistry, University of New Hampshire, Durham, New Hampshire 03824, and \*\*Amgen, Inc., Thousand Oaks, California 91320

We previously demonstrated that P-selectin binds with high affinity to a trace, homodimeric glycoprotein ligand on human myeloid cells. The ligand carries the sialyl Lewis x (sLe<sup>x</sup>) epitope, a limited number of N-linked glycans, and clustered, sialylated O-linked glycans. In this study we demonstrate that the polypeptide component of this ligand is identical to that of P-selectin glycoprotein ligand-1 (PSGL-1), a molecule recently identified by expression cloning from a human myeloid cell cDNA library. We have examined the effects of glycosidases on purified, radioiodinated PSGL-1 from human neutrophils to further characterize the structure and function of the attached oligosaccharides. We found that PSGL-1 had poly-N-acetyllactosamine, only some of which could be removed with endo-β-galactosidase. The majority of the Le<sup>x</sup> and sLe<sup>x</sup> structures were on endo-β-galactosidase-sensitive chains. Peptide:N-glycosidase F (PNGaseF) treatment removed at least two of the three possible N-linked oligosaccharides from PSGL-1. Expression of Le<sup>x</sup> and sLe<sup>x</sup> was not detectably altered by PNGaseF digestion, indicating that these structures were primarily on O-linked poly-N-acetyllactosamine. Endo-β-galactosidase-treated PSGL-1 retained the ability to bind to P-selectin, suggesting that some of the oligosaccharides recognized by P-selectin were either on enzyme-resistant poly-N-acetyllactosamine or on chains which lack poly-N-acetyllactosamine. PNGaseF treatment did not affect the ability of PSGL-1 to bind to P-selectin, demonstrating that the oligosaccharides required for P-selectin recognition are O-linked. PSGL-1 also bound to E-selectin, but with at least 50-fold lower affinity than to P-selectin. These data suggest that PSGL-1 from human neutrophils displays complex, sialylated, and fucosylated O-linked poly-N-acetyllactosamine that promote high affinity binding to P-selectin, but not to E-selectin.

The selectins are a family of three Ca<sup>2+</sup>-dependent membrane-bound lectins that initiate adhesion of leukocytes to

platelets or endothelial cells under the shear forces found in the venular circulation (1–3). L-selectin, expressed on leukocytes, binds to constitutive or inducible ligands on endothelial cells. E-selectin, expressed by cytokine-activated endothelial cells, and P-selectin, expressed by thrombin-activated platelets and endothelial cells, bind to ligands on myeloid cells and subsets of lymphocytes. Although the selectins interact weakly with small sialylated, fucosylated oligosaccharides such as sialyl Lewis x (sLe<sup>x</sup>; NeuAcα2–3Galβ1–4[Fucα1–3]GlcNAc) (4, 5), they bind with higher affinity to glycans displayed on a limited number of glycoproteins (6–11) or proteoglycans (12). The high affinity ligands are potentially important as physiologic mediators of selectin-mediated leukocyte adhesion during inflammation. Thus, understanding the structural basis for high affinity recognition of these glycoconjugates by the selectins has attracted increasing interest.

A subset of the high affinity selectin ligands consists of mucin-like glycoproteins (7, 10, 11, 13). We have previously characterized one of these molecules, a sialomucin ligand for P-selectin that is expressed by human neutrophils and the human promyelocytic HL-60 cell line (6, 13). Binding of P-selectin to the ligand is Ca<sup>2+</sup> dependent and is abolished by treatment of the ligand with sialidase. The ligand is a homodimer with two disulfide-linked subunits with relative molecular masses of ~120,000 as assessed by SDS-PAGE. It carries many unmodified sialic residues as well as the sLe<sup>x</sup> antigen. The ligand contains at least one PNGaseF-sensitive N-linked glycan that P-selectin does not require for recognition. In contrast, it has clustered, sialylated, O-linked oligosaccharides that render the polypeptide backbone sensitive to cleavage by the enzyme O-sialoglycoprotease. Treatment of intact HL-60 cells with this enzyme eliminates the high affinity binding sites for P-selectin (14) and prevents cell adhesion to immobilized P-selectin (13, 15), without affecting the overall surface expression of sLe<sup>x</sup> (13). These data suggest that this sialomucin, which carries only a small portion of the cell surface sLe<sup>x</sup>, corresponds to functionally important, high affinity binding sites for P-selectin on human myeloid cells.

Sako *et al.* (16) recently isolated a cDNA derived from human HL-60 cells that encodes a glycoprotein ligand for P-selectin in COS cells co-transfected with a cDNA for an α1,3/4-fucosyl-

\* This research was supported in part by Grants HL 45510 (to R. P. M. and K. L. M.) and CA 37626 (to R. D. C.) from the United States Public Health Service and by Grant DIR 9002027 (to S. E.) from the National Science Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ Supported by Clinician Scientist Award 900403 from the American Heart Association with funds contributed in part by the AHA Oklahoma Affiliate. To whom correspondence should be addressed: Dept. of Medicine, University of Oklahoma Health Sciences Center, 825 N. E. 13th St., Oklahoma City, OK 73104.

<sup>1</sup> The abbreviations used are: Le<sup>x</sup>, Lewis x; CHO, Chinese hamster ovary; ConA, concanavalin A; HSA, human serum albumin; PBS, phosphate-buffered saline; PAGE, polyacrylamide gel electrophoresis; PNGaseF, peptide:N-glycosidase F; PSGL-1, P-selectin glycoprotein ligand 1; sLe<sup>x</sup>, sialyl Lewis x; tES, truncated soluble E-selectin; tPS, truncated soluble P-selectin; WGA, wheat germ agglutinin; mAbs, monoclonal antibodies; HBSS, Hanks' balanced salt solution; MOPS, 4-morpholinepropanesulfonic acid; HSA, human serum albumin.



transferase. The recombinant glycoprotein also interacts with E-selectin in cell adhesion experiments. The cDNA-derived amino acid sequence of this molecule, termed P-selectin glycoprotein ligand-1 (PSGL-1), predicts a 402-residue type I membrane protein with a Ser/Thr/Pro-rich extracellular domain that contains three potential sites for addition of *N*-linked oligosaccharides as well as a single cysteine that might promote dimerization. Like the P-selectin glycoprotein ligand that we identified in human neutrophils, recombinant PSGL-1 is a homodimer with a limited number of *N*-linked glycans and many *O*-linked oligosaccharides. However, the identity of the polypeptide components of these two molecules has not been formally proven. Furthermore, the degree to which the oligosaccharides on recombinant PSGL-1 resemble those on the native glycoprotein ligand is unknown.

In this study we have demonstrated that antibodies to a peptide derived from PSGL-1 also recognize the human neutrophil glycoprotein ligand, establishing the identities of the polypeptide backbones of both molecules. To ensure consistency in nomenclature, we have also adopted the term PSGL-1 to refer to this glycoprotein. We have studied the effects of various glycosidases on the structure and function of PSGL-1 from human neutrophils. We found that this molecule, unlike recombinant PSGL-1, is largely resistant to treatment with endo- $\alpha$ -*N*-acetylgalactosaminidase, suggesting that it has few simple core 1 Gal $\beta$ 1-3GalNAc disaccharides linked to serine or threonine. Furthermore, human neutrophil PSGL-1 displays sLe<sup>x</sup> and its nonsialylated counterpart, Le<sup>x</sup>, primarily on *O*-linked poly-*N*-acetylglucosamine. This molecule also binds to E-selectin, but with at least 50-fold lower affinity than to P-selectin. These data suggest that PSGL-1 on human neutrophils presents an array of sialylated, fucosylated, *O*-linked poly-*N*-acetylglucosamine that P-selectin preferentially recognizes.

#### EXPERIMENTAL PROCEDURES

**Materials**—Triton X-100 and sialidase from *Arthrobacter ureafaciens* (75 units/mg, EC 3.2.1.18) were obtained from Calbiochem. Iodobeads<sup>TM</sup>, 3M Emphaze<sup>TM</sup> Biosupport Medium, Protein A-Sepharose CL4B, and Brij-58 were from Pierce. Endo- $\beta$ -galactosidase from *Escherichia freundii* (5 units/mg, EC 3.2.1.103) was obtained from V-Labs, Inc. (Covington, LA). *Streptomyces* sp.142  $\alpha$ 1,3/4-*L*-fucosidase (EC 3.2.1.51) was obtained from PanVera Corp. (Madison, WI), and  $\beta$ -*N*-acetylglucosaminidase from Jack Bean (EC 3.2.1.30, 53 units/mg) was from Sigma. Recombinant peptide:*N*-glycosidase F (EC 3.2.2.18, 25,000 units/mg), and endo- $\alpha$ -*N*-acetylgalactosaminidase from *Diplococcus pneumoniae* (EC 3.2.1.97, *O*-glycanase<sup>TM</sup>) were purchased from Genzyme Corp. (Cambridge, MA). Concanavalin A-Sepharose (ConA, 10 mg/ml resin) was from Pharmacia/LKB (Uppsala, Sweden). Wheat germ agglutinin (WGA) agarose (7.6 mg/ml resin) and tomato lectin were from Vector Laboratories, Inc. (Burlingame, VT). Tomato lectin was coupled to cyanogen bromide-activated Sepharose to a density of 2 mg/ml resin in the presence of 7.5 mg/ml chitotriose as described previously (17).

**Antibodies**—The anti-human P-selectin mAbs S12 and G1 were prepared and characterized as described previously (18, 19). The anti-human E-selectin mAbs CL2 and CL37 (20) were kindly provided by C. Wayne Smith (Baylor College of Medicine, Houston, TX). LeuM1 (CD15, IgM) was purchased from Becton-Dickinson & Co. (San Jose, CA). The CSLEX-1 hybridoma (HB 8580) was purchased from the American Type Culture Collection, and the IgM mAb was purified from ascites fluid by boric acid precipitation and gel filtration on Sepharose CL4B in PBS, pH 7.4, as described previously (21). Goat anti-mouse  $\mu$ -chain-specific IgG and MOPC104E (IgM) were purchased from Cappel-Organon Technika (Durham, NC).

**Peptide Antisera Production and Immunoprecipitations**—Peptides corresponding to residues 42–56 (QATEYDYDFLPEC) and 354–376 (CISSLLPDGGEGPSATANGGLSK) of the cDNA-derived amino acid sequence of PSGL-1 (16) were synthesized on an Applied Biosystems model 431 peptide synthesizer. Peptide 42–56 was coupled to maleimide-activated keyhole limpet hemocyanin (Pierce) through the added cysteine (underlined) and injected into New Zealand White rabbits (Montana State University Animal Resources Center). Immune

sera were collected and used to immunoprecipitate <sup>125</sup>I-PSGL-1. Protein A beads (25  $\mu$ l, 50% suspension) were preincubated with 1  $\mu$ l of a 1:4 dilution of either normal or immune rabbit serum for 90 min at 37 °C. The beads were washed once with 0.1 M NaCl, 20 mM MOPS, pH 7.5, 0.1% Triton X-100 (MBS/0.1% Triton). Purified <sup>125</sup>I-PSGL-1 or <sup>125</sup>I-WGA eluate was then incubated with the beads in the presence or absence of 0.2 mg/ml of either the immunizing peptide or PSGL-1 peptide 354–376 which served as a negative control. After a 90-min incubation at 37 °C, the beads were washed five times with MBS/0.1% Triton and then eluted by boiling for 5 min with SDS sample buffer (22). Immunoprecipitates were analyzed by SDS-PAGE (22), followed by autoradiography. All autoradiograms shown represent the entire gel from the stacking gel interface downward.

**Generation of CHO D<sup>-</sup> Cells Stably Expressing a Soluble Form of E-selectin**—An expression plasmid (ELAM-1:BBG 57, obtained from British BioTechnology Products, Ltd.) containing the full-length cDNA for human E-selectin was used as a template in the polymerase chain reaction to modify the 5' and 3' ends of E-selectin for high level production of recombinant soluble E-selectin in CHO D<sup>-</sup> cells. The oligonucleotide pair (5' GTC CCT CTA GAC CAC CAT GAT TGC TTC ACA GTT T + 5' TCC CGG TCG ACT TAG GGA GCT TCA CAG GT) used in the polymerase chain reaction was designed to introduce an *Xba*I site and a perfect Kozak sequence (23) preceding the E-selectin initiator codon, and to introduce a *Sal*I site following the stop codon introduced after amino acid 550 located between the sixth consensus repeat and the transmembrane domain of full-length E-selectin. DNA sequence analysis verified the changes generated by the polymerase chain reaction as well as the integrity of the remainder of the E-selectin cDNA. The truncated E-selectin (tES) gene was then introduced as an *Xba*I-*Sal*I fragment into the mammalian expression vector pDSRa2 (European patent application A20398753) that was modified to include unique *Xba*I and *Sal*I restriction sites. The tES expression vector was transfected into a CHO cell line deficient in dihydrofolate reductase (CHO D<sup>-</sup>) (24), and transfectants were selected in medium lacking hypoxanthine and thymidine (25). An RNase protection assay was used to screen for transfectants that had high levels of E-selectin-specific mRNA (26).

**Protein Purification**—PSGL-1 was purified from human neutrophil membranes prepared as described previously (6), with the following modifications. Neutrophil membranes extracted with 5% Triton X-100 in 0.1 M NaCl, 20 mM MOPS, pH 7.5, 0.02% NaN<sub>3</sub> were applied to WGA-agarose, and bound proteins were eluted with 0.5 M *N*-acetylglucosamine (WGA eluate). After extensive dialysis against 0.1 M NaCl, 20 mM MOPS, pH 7.5, 0.02% NaN<sub>3</sub>, 0.02% Brij-58, 2 mM CaCl<sub>2</sub>, 2 mM MgCl<sub>2</sub> (equilibration buffer), the WGA eluate was loaded on a truncated P-selectin (tPS)-Emphaze column (10 mg lectin/ml resin). The column was extensively washed with equilibration buffer and eluted with 5 mM EDTA. Ligand-containing fractions were pooled and loaded on a Mono Q PC 1.6/5 column equilibrated with 0.1 M NaCl, 20 mM MOPS, pH 7.5, 2 mM EDTA, 0.02% NaN<sub>3</sub>, 0.02% Brij-58 using a SMART<sup>TM</sup> Micro Separation System (Pharmacia/LKB). The column was developed with a 2-ml linear gradient (0.1–1.0 M NaCl) at 50  $\mu$ l/min. Aliquots of the fractions were iodinated with Na<sup>125</sup>I using Iodobeads<sup>TM</sup> according to the instructions of the manufacturer. The purity of the fractions was assessed by SDS-PAGE and autoradiography of the iodinated fractions.

A recombinant soluble form of P-selectin (tPS) truncated after the ninth consensus repeat was immunoaffinity purified from conditioned medium of permanently transfected 293 cells as described previously (14). As determined by sedimentation velocity and equilibrium analysis, tPS is an asymmetric monomer with a molecular weight of 103,600 Da (14). The *E*<sub>280</sub><sup>1%</sup> for tPS is 12.3 (14).

Recombinant soluble truncated E-selectin (tES) was immunoaffinity purified from conditioned medium of CHO D<sup>-</sup> cells stably secreting tES. Cells were cultured in equal parts of Dulbecco's modified Eagle's medium and Ham's F-12 nutrient mixture, supplemented with 10% fetal bovine serum, 10 units/ml penicillin, and 10 mg/ml streptomycin. At confluence, the cells were shifted to serum-free medium, and conditioned medium was harvested every 7 days. Conditioned medium was clarified, sterile-filtered, using two Sartorius filters (10 and 0.2  $\mu$ m) connected in series, and then concentrated 50-fold (Filtron Maximate, 10 kDa molecular mass cut off). The anti-E-selectin mAb H18/7 (27) was coupled to cyanogen bromide-activated Sepharose to a density of 6 mg/ml resin. The H18/7-Sepharose column was equilibrated at 4 °C with PBS, pH 7.5, containing 1 mM CaCl<sub>2</sub> and 1 mM MgCl<sub>2</sub>, and 500 ml of concentrated medium was loaded on the column (*V*<sub>T</sub> = 200 ml) at a flow rate of 10 cm/h. The column was eluted with 0.1 M glycine, pH 2.7, and the fractions were rapidly neutralized by addition of Tris-HCl, pH 7.5, to a final concentration of 0.1 M. The eluate was dialyzed against 25



mm sodium phosphate, pH 7.5, and loaded on a Q-High Performance anion-exchange column ( $V_T = 200$  ml, Pharmacia) equilibrated in the same buffer. The column was developed with a linear gradient of NaCl (0.17–0.22 M). Purity was assessed by SDS-PAGE and by N-terminal sequence analysis. Protein concentration was determined using an  $E_{280}^{1\%} = 12.5$  that was calculated using a previously described method (28). The yield was 250 mg of tES/500 ml of concentrated conditioned medium.

**Sedimentation Equilibrium Analysis of Soluble tES**—Sedimentation equilibrium experiments were conducted using a four-hole titanium rotor, short column kel-f centerpieces, and sapphire windows in a Beckman Optima XLA analytical ultracentrifuge equipped with on-line Rayleigh interference optics. Data were acquired at rotor speeds of 10,000, 15,000, 20,000, and 25,000 revolutions/min at 20 °C using a television-based camera and an on-line acquisition and analysis system (29, 30). Samples were loaded at concentrations of approximately 1.0, 0.5, 0.25, and 0.13 mg/ml in PBS, pH 7.4. Data were collected at intervals after the estimated time to equilibrium and tested for equilibrium by subtracting successive scans (31). The data within the optical window were selected using the program REEDIT (kindly provided by David Yphantis). Data were analyzed to estimate the monomer molecular weight of tES using the NONLIN program (32). Three or more channels of sedimentation equilibrium data obtained at different loading concentration of tES, radial positions, and angular velocities were fit by simultaneous nonlinear least-squares analysis. For molecular weight calculations, the partial specific volume of tES was estimated to be 0.688 ml/g with an inaccuracy of  $\pm 0.03$  ml/g to account for uncertainty in the carbohydrate composition of tES (30).

**Sedimentation Velocity Analysis of Soluble tES**—Sedimentation velocity experiments were conducted in the Beckman Optima XLA analytical ultracentrifuge using Rayleigh interference optics. Experiments were conducted at 60,000 revolutions/min and 20 °C, using a four-hole titanium rotor, two-channel, charcoal-filled, epon centerpieces, and sapphire windows. Sample was loaded at a concentration of 0.5 mg/ml in PBS, pH 7.4. Sedimentation coefficients were determined from the time derivative of the concentration profile as described (33).

**Protein Iodinations**—The mAbs G1 and CL2 (200  $\mu$ g) were iodinated with 400  $\mu$ Ci of carrier-free  $\text{Na}^{125}\text{I}$  (ICN Biomedical, Inc., Costa Mesa, CA) using Iodobeads<sup>TM</sup>. Free  $^{125}\text{I}$  was removed by gel filtration through a Sephadex G-25 column (PD-10, Pharmacia) equilibrated in 0.1 M NaCl, 20 mM MOPS, pH 7.5, 0.1% Triton X-100. The labeled antibodies were then centrifuged for 30 min at  $90,000 \times g$  in a TL-100 ultracentrifuge (Beckman, Palo Alto, CA). The concentration of labeled proteins was determined with a Micro BCA protein assay kit (Pierce) using the respective unlabeled proteins as standards. The specific activity of the labeled antibodies ranged from 0.5 to 1.0  $\mu$ Ci/ $\mu$ g protein.

PSGL-1 ( $\sim 0.5$   $\mu$ g) or WGA eluate ( $\sim 100$   $\mu$ g) was iodinated with 400  $\mu$ Ci of  $\text{Na}^{125}\text{I}$  and desalted as described above. Radioabeled PSGL-1 was concentrated to  $\sim 200$   $\mu$ l using a Centicon-30<sup>TM</sup> device (Pharmacia/LKB), and gel filtered on a Superose 6 PC 3.2/30 column equilibrated with 0.15 M NaCl, 20 mM MOPS, pH 7.5, 0.02%  $\text{NaN}_3$ , 0.1% Brij-58 using a SMART<sup>TM</sup> Micro Separation System. The specific activity of the radio-labeled ligand ranged from 14 to 30  $\mu$ Ci/ $\mu$ g protein. Labeled proteins were routinely  $>95\%$  trichloroacetic acid precipitable.

**Binding of  $^{125}\text{I}$ -PSGL-1 to Immobilized tPS or tES**—tPS or tES (50  $\mu$ l, 10  $\mu$ g/ml) in HBSS, 0.02%  $\text{NaN}_3$  (HBSS/Az) were incubated overnight at 4 °C in microtiter plates (Immulon I Removawell<sup>TM</sup> Strips, Dynatech Laboratories, Inc., Chantilly, VA). After three washes with HBSS/Az the wells were blocked with HBSS/Az, 1% human serum albumin (HSA) for 2 h at 22 °C. The wells were then washed once with HBSS/Az and  $^{125}\text{I}$ -PSGL-1 diluted in HBSS/Az, 0.1% HSA was added (50  $\mu$ l, 5,000–10,000 counts/min/well). After 1 h at 22 °C, the wells were rapidly washed five times with HBSS/Az using a 12-channel plate washer (Nunc-Immuno Wash 12, Nunc Inc., Naperville, IL). The individual wells were then counted in a  $\gamma$  counter. All assays were performed in quadruplicate. In certain assays binding of  $^{125}\text{I}$ -PSGL-1 to immobilized tPS was measured in the presence of increasing concentrations of fluid-phase tPS or tES.

**Binding of  $^{125}\text{I}$ -PSGL-1 to Immobilized Anti-carbohydrate Antibodies**—Goat anti-mouse  $\mu$ -chain specific IgG (50  $\mu$ l, 10  $\mu$ g/ml) in HBSS/Az was incubated overnight at 4 °C in Immulon I Removawell<sup>TM</sup> strips. The wells were washed and blocked as described above. Then, either CSLEX-1, LeuM1, or MOPC104E was added (50  $\mu$ l, 5  $\mu$ g/ml). After incubation for 1 h at 22 °C, the wells were washed three times with HBSS/Az, and  $^{125}\text{I}$ -PSGL-1 in HBSS/Az, 0.1% HSA (50  $\mu$ l, 5,000–10,000 counts/min/well) was added. After 1 h the wells were washed as described above, and individual wells were counted in a  $\gamma$  counter. All assays were performed in quadruplicate.

**Enzyme Digestion of  $^{125}\text{I}$ -PSGL-1**—Digestions with sialidase (0.2 unit/ml, 15 h), endo- $\beta$ -galactosidase (2 unit/ml, 15 h), and  $\alpha$ 1,3/4-L-fucosidase (0.32 milliunit/ml, 15 h) were performed in 0.1 M NaCl, 50 mM sodium acetate, pH 5.5. Digestions with PNGaseF (40 units/ml, 15 h) were done in 0.1 M sodium phosphate, pH 8.6. PSGL-1 was generally not pretreated with SDS because PNGase F had similar effects with or without SDS pretreatment. Sequential digestions with sialidase (0.2 unit/ml, 2 h) followed by endo- $\alpha$ -N-acetylgalactosaminidase (0.1 unit/ml, 15 h) were performed in 0.1 M NaCl, 50 mM sodium acetate, pH 6.0. All glycosidase digestions were performed at 37 °C in the presence of 20  $\mu$ M leupeptin, 30  $\mu$ M antipain, 1 mM benzamide, and 0.02%  $\text{NaN}_3$ .

**Lectin Affinity Chromatography**— $^{125}\text{I}$ -PSGL-1 was applied to a concanavalin A (ConA)-Sepharose column ( $V_T = 1$  ml) equilibrated in 0.1 M NaCl, 20 mM Tris, pH 7.5, 2 mM  $\text{CaCl}_2$ , 0.1% Brij-58, 0.02%  $\text{NaN}_3$ . The column was washed with five column volumes of equilibration buffer and eluted with 0.5 M  $\alpha$ -methylmannoside in equilibration buffer preheated to 65 °C. Tomato lectin chromatography was similarly performed except that the column ( $V_T = 1$  ml) was equilibrated with 0.1 M NaCl, 20 mM MOPS, pH 7.5, 0.1% Brij-58, 0.02%  $\text{NaN}_3$ , and step eluted with 20 mg/ml chitotriose in equilibration buffer. Greater than 85% of the counts loaded on the lectin columns were recovered.

**Site Density Determinations**—mAbs to P-selectin (G1) or E-selectin (CL2) that inhibit interactions with myeloid cells were used for site density measurements. Wells were coated with tPS or tES (50  $\mu$ l, 10  $\mu$ g/ml) and blocked as described above. Saturating concentrations of  $^{125}\text{I}$ -labeled mAb (2  $\mu$ g/ml) in HBSS/Az, 0.1% HSA were added to the wells and incubated for 1 h at 22 °C. The wells were washed five times and then counted in a  $\gamma$  counter. Specific binding was defined as counts/min bound to selectin-coated wells minus counts/min bound to HSA-coated wells. Site densities were calculated assuming monovalent binding of antibody at saturation. All assays were performed in quadruplicate.

**Neutrophil Adhesion Assay**—Human neutrophils were isolated from heparinized blood from consenting volunteer donors by dextran sedimentation, hypotonic lysis, and Ficoll-Hypaque density gradient centrifugation as described previously (34). Quantitation of neutrophil adhesion to tPS or tES (50  $\mu$ l, 10  $\mu$ g/ml), immobilized on Immulon I Removawell<sup>TM</sup> microtiter plates, was performed as described previously (19).

## RESULTS

**An Antiserum to a PSGL-1 Peptide Recognizes the P-selectin Glycoprotein Ligand from Human Neutrophils**—We wished to determine whether the polypeptide core of the sialomucin ligand for P-selectin that we characterized in human neutrophils and HL-60 cells (6, 13) was immunologically related to PSGL-1, a glycoprotein ligand for P-selectin identified by expression cloning from a HL-60 cDNA library (16). We prepared antiserum in rabbits against a peptide corresponding to residues 42–56 of the cDNA-derived amino acid sequence of PSGL-1 and tested whether this antiserum could recognize the glycoprotein ligand that we isolated by P-selectin affinity chromatography. Fig. 1A demonstrates the radiochemical purity of the  $^{125}\text{I}$ -P-selectin glycoprotein ligand used in this study. It exhibited a relative molecular mass of  $\sim 250,000$  under nonreducing conditions and  $\sim 120,000$  under reducing conditions. Thus, the radioiodinated glycoprotein contained two disulfide-linked subunits of  $M_r$  120,000, consistent with previous results using protein blotting or metabolic labeling (6). As noted previously, a small portion of the dimeric glycoprotein was resistant to reduction under the conditions used. Fig. 1B shows that the immune serum, but not normal rabbit serum, precipitated the purified  $^{125}\text{I}$ -P-selectin ligand. Precipitation by the antiserum was specific for the 42–56 peptide sequence because it was completely inhibited by inclusion of this peptide, but not by a control peptide corresponding to amino acids 354–376 of PSGL-1. The specificity of the antiserum was further demonstrated by its selective immunoprecipitation of the ligand from a crude mixture of radiolabeled neutrophil membrane proteins that were eluted from a WGA column (Fig. 1C, arrows).

We also sequenced two tryptic peptides from the P-selectin glycoprotein ligand purified from human neutrophils. The se-

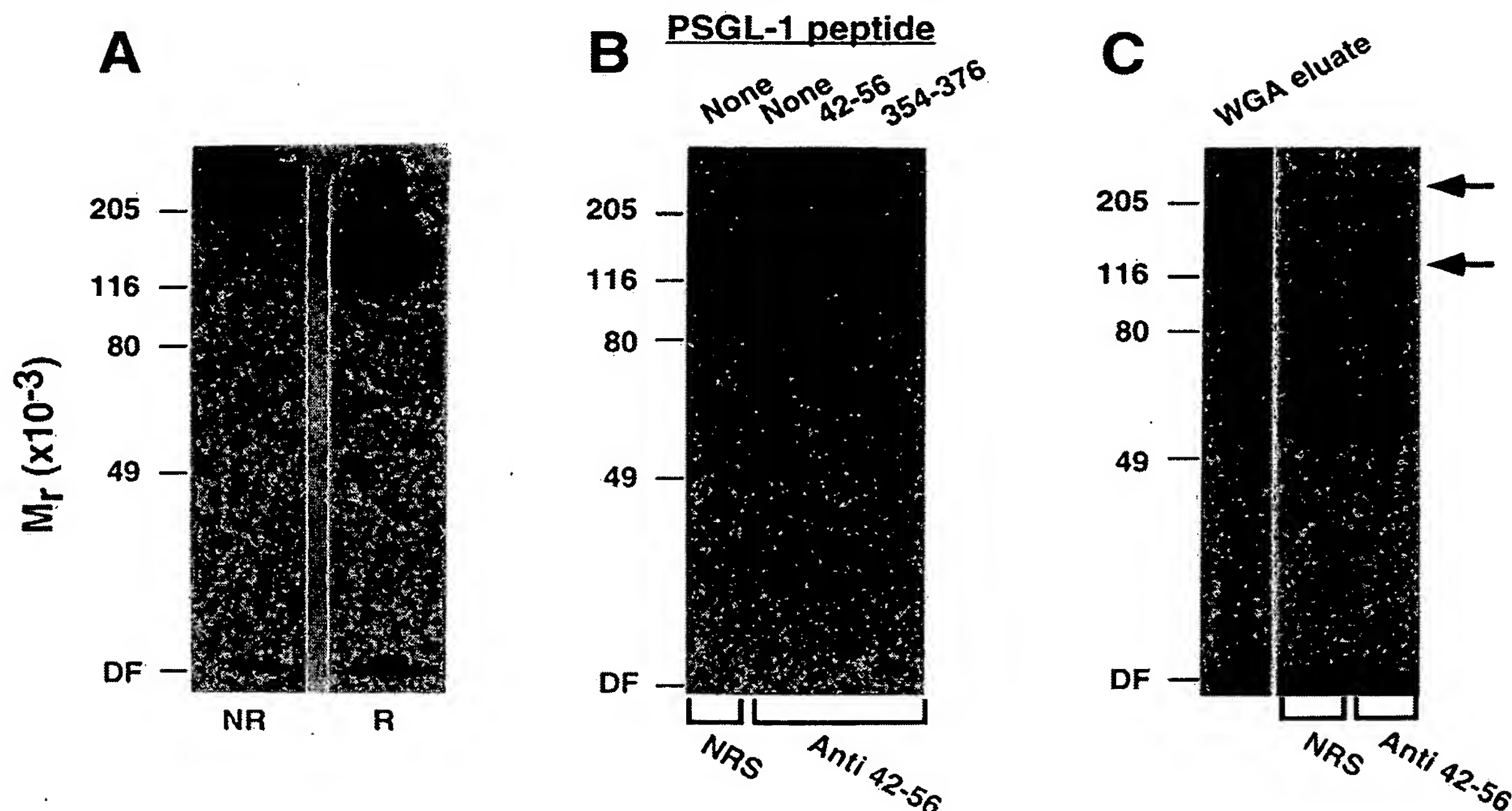


FIG. 1. Immunoprecipitation of the P-selectin glycoprotein ligand by anti-PSGL-1 peptide antibodies. A, PSGL-1 isolated from human neutrophils was radiolabeled as described under "Experimental Procedures" and electrophoresed in an SDS, 7.5% polyacrylamide gel under nonreducing (NR) and reducing (R) conditions. B,  $^{125}\text{I}$ -PSGL-1 was immunoprecipitated with either normal rabbit serum (NRS) or rabbit antiserum directed against a peptide corresponding to residues 42–56 of PSGL-1 (*anti 42–56*) in the presence or absence of 0.2 mg/ml of the immunizing peptide or of a control peptide corresponding to residues 354–376 of PSGL-1. C,  $^{125}\text{I}$ -WGA eluate from human neutrophil membrane extracts was immunoprecipitated using either normal rabbit serum or anti 42–56 serum. The samples in B and C were electrophoresed in SDS, 7.5% polyacrylamide gels under reducing conditions. All gels were then analyzed by autoradiography. DF = dye front.

quences of these peptides, His-Met-Tyr-Pro-Val-Arg and Pro-Gly-Lys-Thr-Pro-Glu-Pro, are identical to amino acids 340–345 and 380–386 in the cDNA-derived sequence of PSGL-1 (data not shown). These data and the results from the immunoprecipitation experiments establish the identity of the polypeptide backbones of the P-selectin ligands studied by both groups. To ensure consistency in nomenclature, we will use the term PSGL-1 to refer to this polypeptide, although the possibility exists that it is differentially glycosylated in various cell types.

**Purified  $^{125}\text{I}$ -PSGL-1 Binds Specifically to P-selectin**— $^{125}\text{I}$ -PSGL-1 isolated from human neutrophils bound in a time-dependent fashion to tPS immobilized on microtiter wells (Fig. 2A). Binding increased as a function of the amount of tPS coated on the wells (Fig. 2B). Binding was  $\text{Ca}^{2+}$  dependent and was abolished by G1, a mAb to P-selectin that inhibits leukocyte adhesion to P-selectin, but not by S12, a mAb that does not block adhesion (Fig. 2B). In addition,  $84.6 \pm 2.4\%$  (mean  $\pm$  S.D.,  $n = 4$ ) of the radiolabeled glycoprotein rebound to a column of recombinant, soluble P-selectin (tPS) and was eluted with buffer containing EDTA. This result demonstrates that the function of  $^{125}\text{I}$ -PSGL-1 was not substantially altered by iodination.

**PSGL-1 Contains Substituted Poly-N-acetyllactosamine**—PSGL-1 from human myeloid cells contains clustered, sialylated O-linked oligosaccharides that are released from the polypeptide backbone by  $\beta$ -elimination (13). To determine whether these O-linked glycans had simple structures, we treated  $^{125}\text{I}$ -PSGL-1 with sialidase, then with endo- $\alpha$ -N-acetylgalactosaminidase, an enzyme that releases non-sialylated Gal $\beta$ 1–3GalNAc core 1 disaccharides, but not more complex O-linked glycans, from serine or threonine residues (35). Fig. 3A shows that treatment with sialidase slowed the electrophoretic mobility of PSGL-1, confirming previous results (6). Subsequent addition of endo- $\alpha$ -N-acetylgalactosaminidase increased the mobility of PSGL-1 only slightly more than sialidase-treated PSGL-1, suggesting that very few of the O-linked

oligosaccharides were simple structures that were susceptible to this enzyme.

PSGL-1 is fucosylated because it contains the sialylated, fucosylated tetrasaccharide antigen sLe<sup>x</sup> (13). Treatment of the glycoprotein with  $\alpha$ 1,3/4-fucosidase, which removes fucose attached to a penultimate GlcNAc, had no effect on electrophoretic mobility (Fig. 3B), although it did remove fucose residues that were part of the Le<sup>x</sup> epitopes on the molecule (see Fig. 5). We next treated the ligand with endo- $\beta$ -galactosidase, which hydrolyzes internal  $\beta$ 1–4 linkages between galactose and N-acetylglucosamine (Gal $\beta$ 1–4GlcNAc) in extended unbranched poly-N-acetyllactosamine chains (36, 37). This enzyme slightly accelerated the mobility and lessened the electrophoretic heterogeneity of PSGL-1, suggesting that PSGL-1 contained some poly-N-acetyllactosamine (Fig. 3B, see also Fig. 4C).

To confirm the presence of poly-N-acetyllactosamine on PSGL-1, we tested the ability of  $^{125}\text{I}$ -PSGL-1 to bind to a column containing immobilized tomato lectin, which avidly binds to poly-N-acetyllactosamine. A single poly-N-acetyllactosamine chain is sufficient to confer binding of a glycoprotein to tomato lectin in this system (17). Greater than 95% of the PSGL-1 bound to the column, indicating that essentially every molecule contained at least one poly-N-acetyllactosamine (Table I). Treatment of PSGL-1 with endo- $\beta$ -galactosidase modestly reduced to 84% the amount of material binding to the column. This suggests that most if not all of the poly-N-acetyllactosamine could be removed from a portion of the PSGL-1 molecules by endo- $\beta$ -galactosidase. Because substitutions of poly-N-acetyllactosamine with sialic acid and/or fucose inhibit the efficiency of this enzyme, we pretreated PSGL-1 with sialidase and fucosidase before addition of endo- $\beta$ -galactosidase. Treatment with sialidase and fucosidase had no effect on binding of PSGL-1 to the column, consistent with the fact that substitutions with sialic acid or fucose do not affect binding of poly-N-acetyllactosamine to tomato lectin. However, subsequent addi-



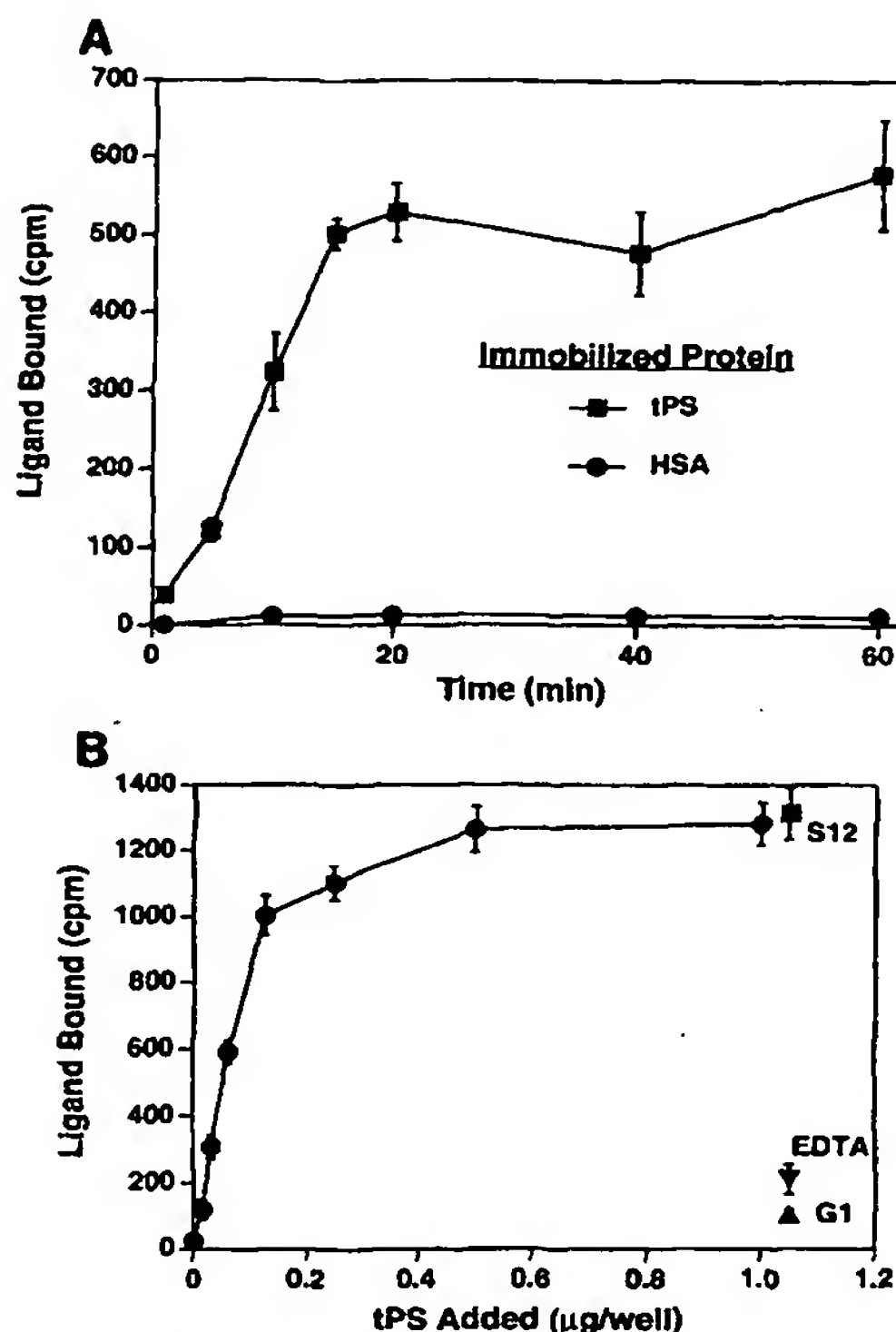


FIG. 2. Binding of <sup>125</sup>I-PSGL-1 to immobilized tPS. A, <sup>125</sup>I-PSGL-1 was incubated for varying times in microtiter wells coated with tPS (50  $\mu$ l, 10  $\mu$ g/ml) and then blocked with HSA, or in control wells without tPS that were blocked with HSA. B, <sup>125</sup>I-PSGL-1 was incubated for 1 h in microtiter wells in which varying amounts of tPS (50  $\mu$ l, 0–20  $\mu$ g/ml) had been immobilized. As indicated, some incubations were performed in the presence of 8 mM EDTA or 40  $\mu$ g/ml of the mAbs S12 or G1 to P-selectin. The data represent the mean  $\pm$  S.D. of quadruplicate wells and are representative of three independent experiments.

tion of endo- $\beta$ -galactosidase reduced to 69% the portion of the glycoprotein binding to the column. Because tomato lectin binds weakly to terminal GlcNAc residues that are exposed by endo- $\beta$ -galactosidase treatment (38), we then added  $\beta$ -N-acetylglucosaminidase to remove these terminal residues. Only 53% of this material bound to tomato lectin. These data clearly demonstrated that PSGL-1 contained poly-N-acetylglucosamine. Furthermore, at least some of these chains were resistant to endo- $\beta$ -galactosidase because of substitutions with sialic acid and/or fucose. The inability of sialidase, fucosidase, endo- $\beta$ -galactosidase, and  $\beta$ -N-acetylglucosaminidase to eliminate binding of all PSGL-1 molecules to tomato lectin may reflect enzyme resistance because of the clustered nature of O-linked glycans, the presence of internal fucose residues, and/or the presence of additional substitutions such as sulfate on the poly-N-acetylglucosamine.

**PSGL-1 Contains a Limited Number of Heterogeneous N-Linked Oligosaccharides**—Fig. 4A shows that PNGaseF treatment slightly increased the electrophoretic mobility of <sup>125</sup>I-PSGL-1, consistent with previous studies indicating that all molecules contained a limited number of N-linked glycans sensitive to this enzyme (6). The cDNA-derived sequence of PSGL-1 indicates that the molecule contains only three potential sites for attachment of N-linked oligosaccharides (16). To explore possible heterogeneity of the N-linked structures, we applied sham-treated and PNGaseF-treated <sup>125</sup>I-PSGL-1 to a column of ConA, a plant lectin that binds avidly to high mannose N-linked glycans, less avidly to complex biantennary and hybrid N-linked oligosaccharides, and very weakly to complex triantennary and tetraantennary N-linked chains (39). We

found that  $41.9 \pm 4.2\%$  of the sham-treated ligand bound to ConA, whereas only  $3.7 \pm 1.0\%$  of the PNGaseF-treated material bound (mean  $\pm$  S.D.,  $n = 4$ ). Greater than 85% of the counts loaded on the column were recovered. This result demonstrated that PNGaseF removed N-linked glycans recognized by ConA from PSGL-1.

Fig. 4B shows that the fraction of PSGL-1 not bound by ConA migrated slightly slower than the fraction bound by ConA. To confirm that the ConA-unbound material also contained N-linked oligosaccharides, we treated both the bound and unbound fractions with PNGaseF (Fig. 4C). The enzyme decreased the apparent molecular masses of the ConA-unbound and ConA-bound fractions by 14.5 and 13.7 kDa, respectively. Thus, both fractions had PNGaseF-sensitive N-linked glycans. These data suggested that the ConA-unbound fraction contained triantennary and/or tetraantennary N-linked glycans that ConA did not recognize. The largest N-linked glycan to be described on myeloid cells (molecular mass  $\sim 6,600$ ) is a disialylated tetraantennary structure with poly-N-acetylglucosamine sequences on each antenna and both external and core fucose (40). Therefore, the observed change in electrophoretic mobility after PNGaseF treatment is consistent with the removal of at least two and possibly three complex N-linked glycans. Even after PNGaseF treatment, the ConA-unbound material migrated slightly slower than the ConA-bound material. This may indicate that the ConA-unbound fraction retained a single N-linked chain that was resistant to PNGaseF. Alternatively, there may be other structural differences between the ConA-bound and -unbound fractions, of which the most likely is heterogeneity in O-linked glycosylation.

Treatment with sialidase or endo- $\beta$ -galactosidase had similar effects on the electrophoretic mobilities of the ConA-unbound and -bound fractions, probably because these enzymes primarily affected the abundant O-linked oligosaccharides (Fig. 4C). Endo- $\beta$ -galactosidase produced similar increases in the electrophoretic mobilities of PSGL-1 pretreated with PNGaseF (data not shown) and PSGL-1 that contained all its N-linked oligosaccharides (Figs. 3B and 4C), suggesting that at least some of the poly-N-acetylglucosamine was on the O-linked glycans.

**PSGL-1 Expresses Le<sup>x</sup> and sLe<sup>x</sup> on O-Linked Poly-N-acetylglucosamine**—To determine whether PSGL-1 expresses Le<sup>x</sup> or sLe<sup>x</sup> on poly-N-acetylglucosamine, we developed an assay to measure the binding of <sup>125</sup>I-PSGL-1 to an immobilized monoclonal IgM antibody to sLe<sup>x</sup> (CSLEX-1) or Le<sup>x</sup> (LeuM1), or to an immobilized irrelevant IgM mAb (MOPC104E). Fig. 5 demonstrates that PSGL-1 bound to CSLEX-1 but not to MOPC104E, confirming previous immunoblotting data that the glycoprotein expressed sLe<sup>x</sup> (13). Greater than 85% of the <sup>125</sup>I-PSGL-1 molecules bound to a CSLEX-1 affinity column, indicating that virtually all molecules contained sLe<sup>x</sup> (data not shown). PSGL-1 also bound to LeuM1, indicating that it contained Le<sup>x</sup>, although it is possible that desialylation generated this structure during purification of the glycoprotein (Fig. 5). Treatment of PSGL-1 with sialidase eliminated binding to CSLEX-1 and increased binding to LeuM1, consistent with the known specificities of these antibodies. Treatment of PSGL-1 with fucosidase had no effect on binding to CSLEX-1, in keeping with the inhibition of this enzyme by terminal sialic acid (41, 42). In contrast, the fucosidase eliminated the binding of PSGL-1 to LeuM1.

After establishing the specificity of the assay, we examined whether treatment of PSGL-1 with endo- $\beta$ -galactosidase affected the expression of Le<sup>x</sup> or sLe<sup>x</sup> (Fig. 5). This enzyme abolished binding of PSGL-1 to immobilized LeuM1 and decreased binding to CSLEX-1 by  $64 \pm 14\%$  (mean  $\pm$  S.D.,  $n = 4$ ). These



FIG. 3. SDS-PAGE of glycosidase-treated  $^{125}\text{I}$ -PSGL-1.  $^{125}\text{I}$ -PSGL-1 was treated with the indicated glycosidases, then electrophoresed in SDS, 7.5% polyacrylamide gels under reducing conditions. The gels were then analyzed by autoradiography. The results are representative of four independent experiments. *Endo- $\alpha$ -GalNAcase* = endo- $\alpha$ -N-acetyl-galactosaminidase; *DF* = dye front.

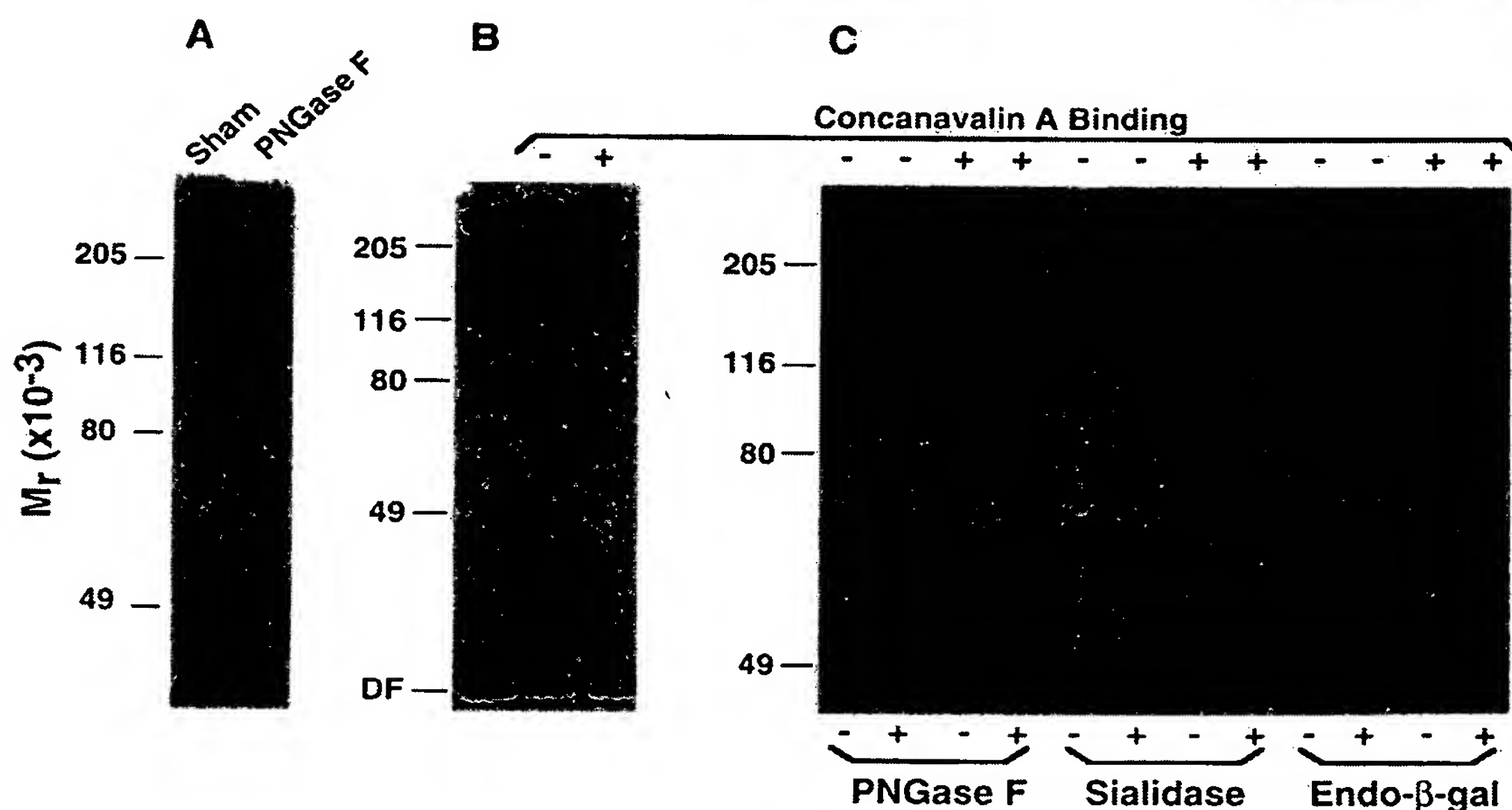
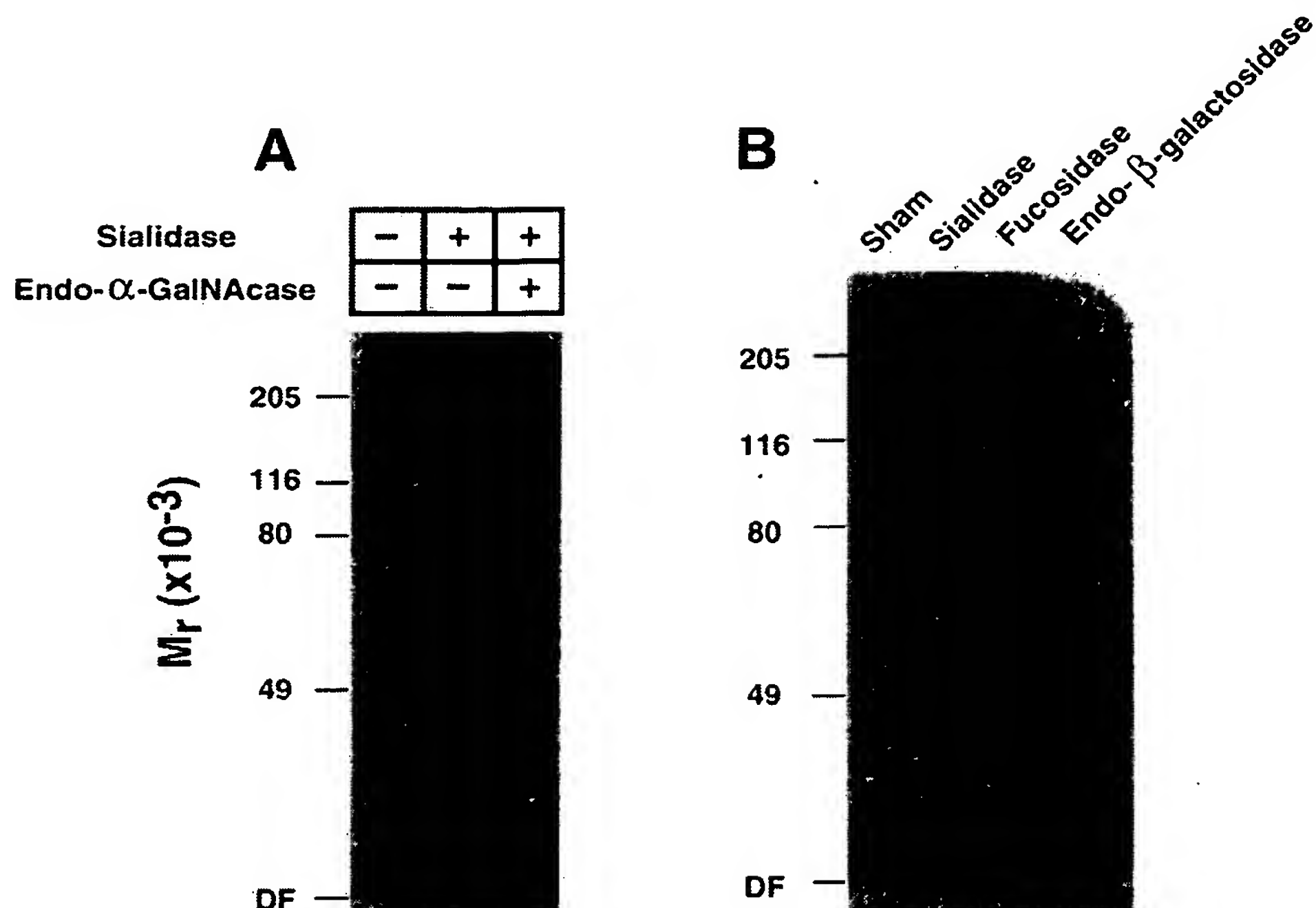


FIG. 4. Effects of glycosidase treatment on ConA-fractionated  $^{125}\text{I}$ -PSGL-1. A,  $^{125}\text{I}$ -PSGL-1 was sham-treated or treated with PNGaseF, then electrophoresed in an SDS, 7.5% polyacrylamide gel under reducing conditions and analyzed by autoradiography. B,  $^{125}\text{I}$ -PSGL-1 was fractionated on ConA-Sepharose as described under "Experimental Procedures." Unfractionated, ConA-bound, and ConA-unbound samples were electrophoresed in an SDS, 7.5% polyacrylamide gel under reducing conditions and analyzed by autoradiography. C, ConA-unbound and ConA-bound  $^{125}\text{I}$ -PSGL-1 was sham-treated or treated with the indicated glycosidase, then electrophoresed in an SDS, 7.5% polyacrylamide gel under reducing conditions and analyzed by autoradiography. *DF* = dye front.

data clearly indicated that poly-*N*-acetylglucosamine cleaved by endo- $\beta$ -galactosidase carried some, and perhaps all, of the  $\text{Le}^x$  as well as a portion of the  $\text{sLe}^x$  on PSGL-1. The remaining  $\text{sLe}^x$  structures may be on branched or substituted poly-*N*-acetylglucosamine that was resistant to the enzyme or on chains that lacked poly-*N*-acetylglucosamine.

In contrast to the effects of endo- $\beta$ -galactosidase, PNGaseF did not affect binding of PSGL-1 to immobilized LeuM1 or CSLEX-1, indicating that few, if any,  $\text{Le}^x$  or  $\text{sLe}^x$  structures were present on PNGaseF-sensitive *N*-linked oligosaccharides (Fig. 5). In conjunction with the other data in Fig. 5, this observation indicated that *O*-linked oligosaccharides carried most, and perhaps all, of the  $\text{Le}^x$  and  $\text{sLe}^x$  structures. Further-

more, at least some of these structures were on *O*-linked poly-*N*-acetylglucosamine.

**Effects of Glycosidases on Binding of PSGL-1 to P-selectin**—We next examined the effects of the glycosidases on binding of  $^{125}\text{I}$ -PSGL-1 to immobilized tPS. As shown in Fig. 6, treatment with sialidase, which eliminated binding to the anti- $\text{Le}^x$  antibody but enhanced binding to the anti- $\text{Le}^x$  antibody, completely prevented binding of PSGL-1 to tPS. In contrast, treatment with fucosidase, which eliminated binding to the anti- $\text{Le}^x$  antibody but not to the anti- $\text{sLe}^x$  antibody, had no effect on the interaction of PSGL-1 with tPS. These data are consistent with previous results from other assays that high affinity binding of P-selectin to myeloid cells or to PSGL-1

TABLE I  
Effects of glycosidases on binding of  $^{125}\text{I}$ -PSGL-1  
to a tomato lectin affinity column

$^{125}\text{I}$ -PSGL-1 was treated with the indicated glycosidase(s) and applied to a column of immobilized tomato lectin ( $V_T = 1$  ml, 2 mg lectin/ml resin). After washing, the column was eluted with 20 mg/ml of chitotriose. The percentage of PSGL-1 bound was defined as the radioactivity eluted by chitotriose divided by the sum of the radioactivity in the column flowthrough, wash, and eluate.

Enzyme treatment	Percentage bound
Sham	>95
Endo- $\beta$ -galactosidase	84
Sialidase + fucosidase	>95
Sialidase + fucosidase + Endo- $\beta$ -galactosidase	69
Sialidase + fucosidase + Endo- $\beta$ -galactosidase + $\beta$ -N-acetylglucosaminidase	53
Peptide: N-glycosidase F	>95

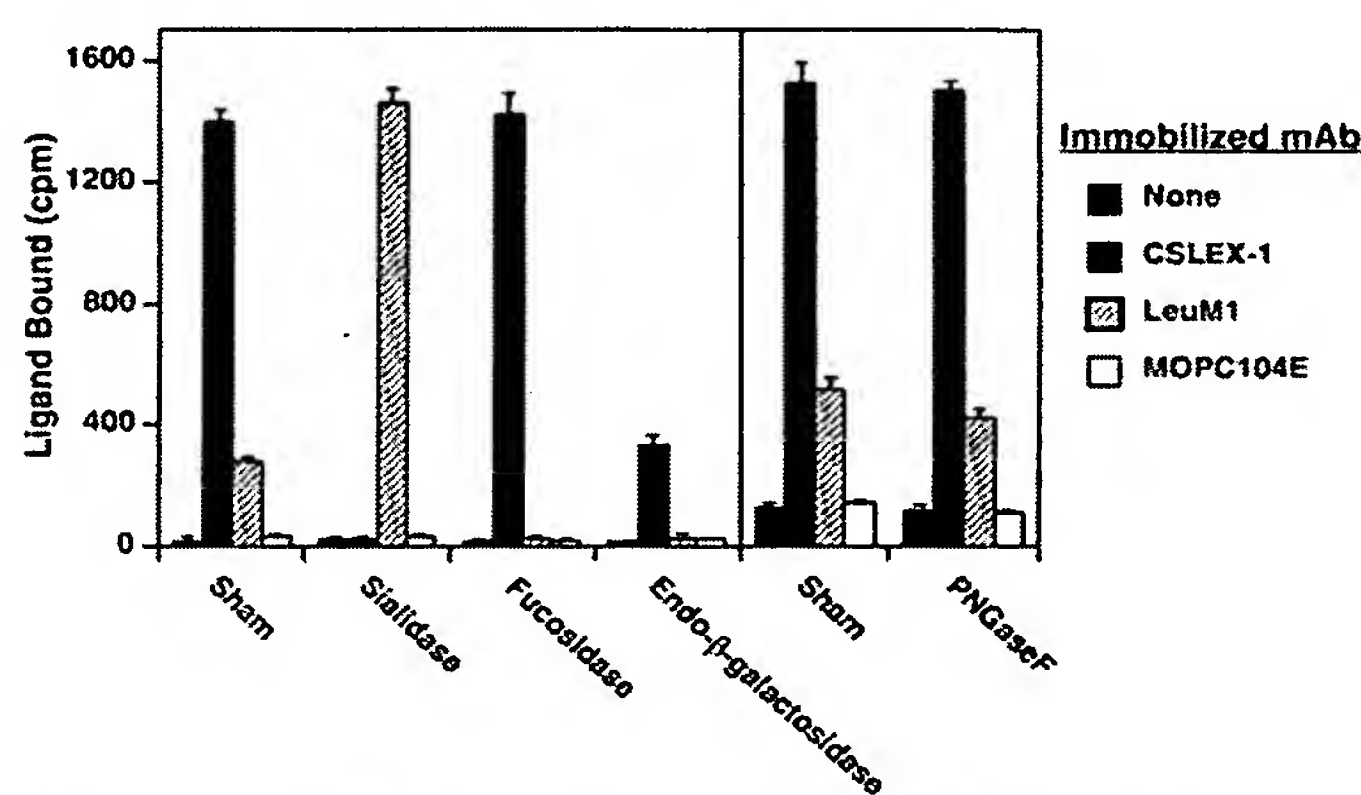


FIG. 5. Effects of glycosidases on binding of  $^{125}\text{I}$ -PSGL-1 to immobilized anti-carbohydrate monoclonal antibodies.  $^{125}\text{I}$ -PSGL-1 was sham-treated or treated with the indicated glycosidase. Samples were analyzed for their ability to bind to immobilized anti-sLe<sup>x</sup> (CSLEX-1), anti-Le<sup>x</sup> (LeuM1), or an irrelevant control IgM mAb, MOPC104E. The data represent the mean  $\pm$  S.D. of quadruplicate wells and are representative of three independent experiments.

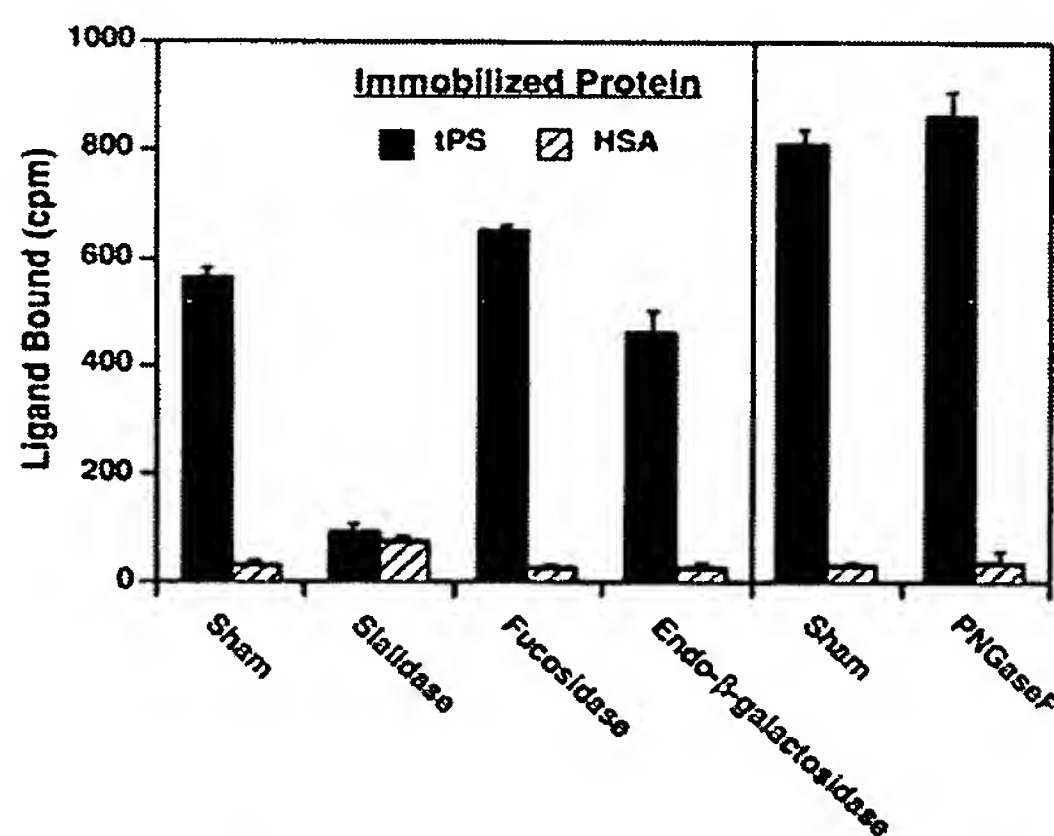


FIG. 6. Effects of glycosidases on binding of  $^{125}\text{I}$ -PSGL-1 to immobilized tPS.  $^{125}\text{I}$ -PSGL-1 was sham-treated or treated with the indicated glycosidase. Samples were analyzed for their ability to bind to microtiter wells coated with tPS (50  $\mu\text{l}$ , 10  $\mu\text{g}/\text{ml}$ ) and then blocked with HSA, or to control wells lacking tPS that were blocked with HSA. The data represent the mean  $\pm$  S.D. of quadruplicate wells and are representative of three independent experiments.

required sialic acid on the ligand (6, 43, 44).

Treatment of PSGL-1 with endo- $\beta$ -galactosidase resulted in only a modest reduction in binding to immobilized tPS (Fig. 6). In four independent experiments, endo- $\beta$ -galactosidase re-

duced binding by  $25 \pm 11\%$  (mean  $\pm$  S.D.,  $n = 4$ ). These data suggested that some, but not all, of the sialylated chains required for P-selectin recognition were on poly-*N*-acetylglucosamine that were sensitive to the enzyme. The remaining glycans required for recognition might be on poly-*N*-acetylglucosamine that were resistant to endo- $\beta$ -galactosidase or on oligosaccharides that lacked poly-*N*-acetylglucosamine. Endo- $\beta$ -galactosidase inhibited binding of PSGL-1 to the anti-sLe<sup>x</sup> antibody more than it inhibited binding to tPS. However, we could not accurately correlate the levels of sLe<sup>x</sup> with the interaction with tPS, since we do not know how many sLe<sup>x</sup> structures were required for PSGL-1 to bind to either the antibody or to tPS in these assays.

Treatment of PSGL-1 with PNGaseF did not affect its ability to bind to tPS (Fig. 6), supporting the conclusion that *N*-linked oligosaccharide played little if any role in P-selectin recognition.

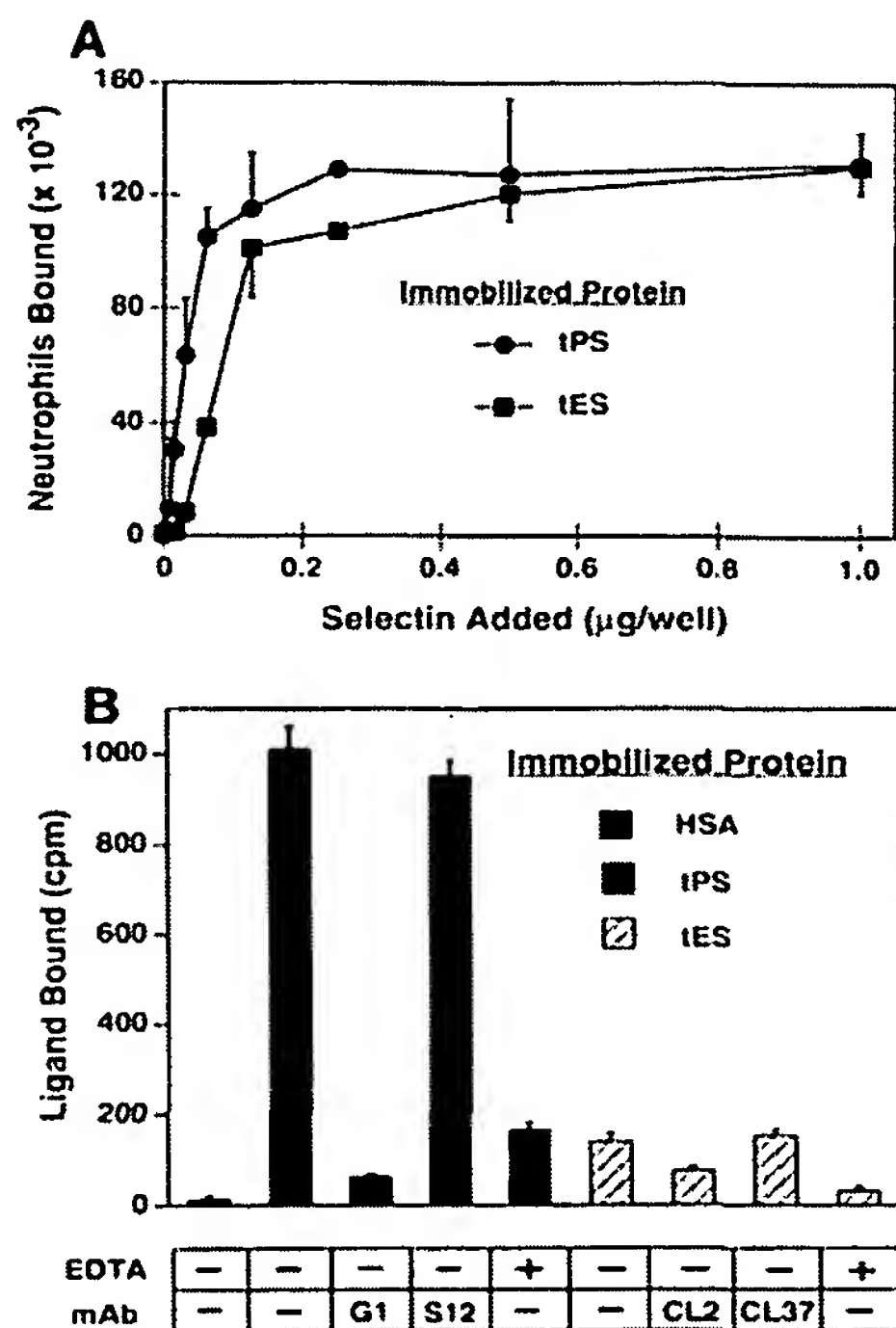
**PSGL-1 Binds to E-selectin But with Lower Affinity Than to P-selectin**—Both P- and E-selectin bind to human leukocytes and interact weakly with the tetrasaccharide, sLe<sup>x</sup> (44–48). Since PSGL-1 expresses sLe<sup>x</sup>, we examined whether PSGL-1 could also bind to E-selectin. We expressed a recombinant soluble form of E-selectin, termed tES, that was truncated immediately before the transmembrane domain. Sedimentation equilibrium measurements indicated that tES, like tPS (14), was monomeric, with no mass action association over the protein concentrations examined. We determined the molecular mass of tES to be 85,800 Da, of which 58,400 Da was derived from the polypeptide backbone and the remainder from attached *N*-linked oligosaccharides. Sedimentation velocity measurements indicated that tES was highly asymmetric. Assuming the molecule to be a hydrated prolate ellipsoid, it had a sedimentation coefficient of 3.60 s and an axial ratio of 19.7. Thus tES, like tPS (14), was a rigid, elongated monomer. The availability of soluble, monomeric forms of both selectins allowed us to directly compare their binding affinities for PSGL-1.

Fig. 7A demonstrates that immobilized tES, like tPS, supported adhesion of neutrophils, confirming its functional status. The degree of adhesion increased in proportion to the coating concentration of each selectin. Neutrophil adhesion to immobilized tES was  $\text{Ca}^{2+}$  dependent and was inhibited by CL2, a mAb to E-selectin that blocks adhesion of myeloid cells to E-selectin, but not by CL37, a mAb to E-selectin that does not block adhesion (20) (data not shown).

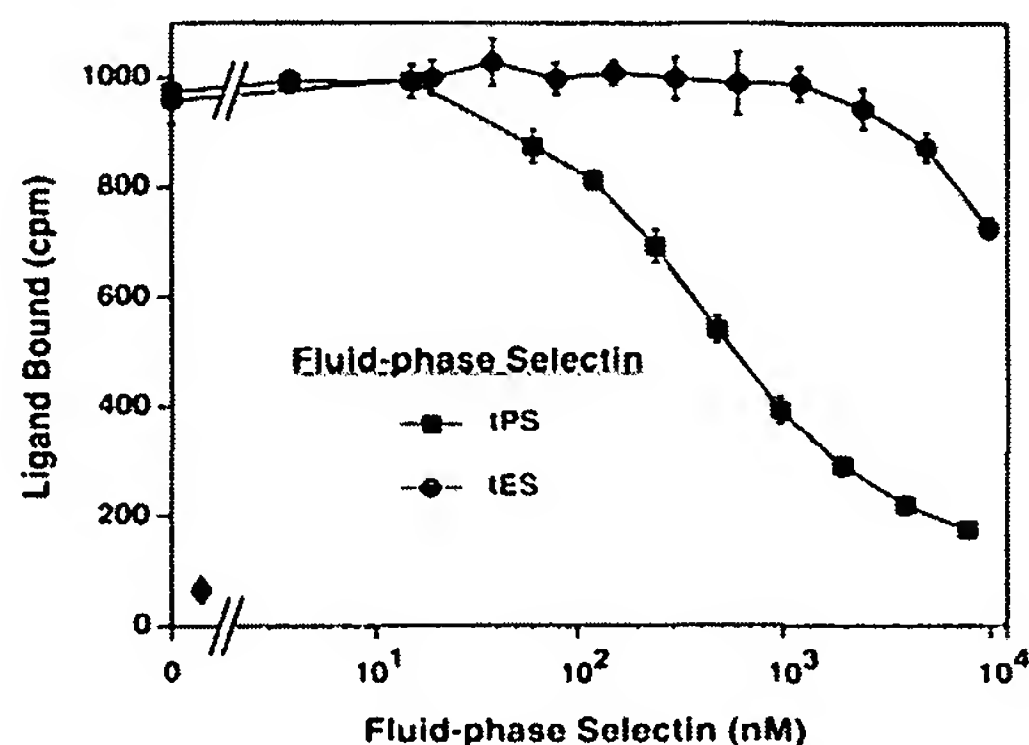
To examine the relative abilities of immobilized tPS and tES to interact with PSGL-1, we measured binding of  $^{125}\text{I}$ -PSGL-1 to each selectin coated at a concentration that promoted maximal neutrophil adhesion. Measurements using  $^{125}\text{I}$ -labeled mAbs indicated that the site densities of the immobilized selectins were identical ( $558 \pm 29$  and  $561 \pm 16$  sites/ $\mu\text{m}^2$  for tPS and tES, respectively (mean  $\pm$  S.D.,  $n = 3$ )). Fig. 7B shows that  $^{125}\text{I}$ -PSGL-1 bound to both tPS and tES, although binding was much higher to tPS than to tES. In each case, binding was specific, as it was blocked by EDTA and by an inhibitory mAb to the selectin. The data from this nonequilibrium binding assay suggested that E-selectin also recognized PSGL-1, but less well than did P-selectin.

To examine the relative affinities of tPS and tES for PSGL-1 under equilibrium conditions, we measured the ability of fluid-phase tPS or tES to inhibit binding of  $^{125}\text{I}$ -PSGL-1 to immobilized tPS. Fig. 8 shows that fluid-phase tPS inhibited binding of  $^{125}\text{I}$ -PSGL-1 to immobilized tPS with an  $\text{IC}_{50}$  of  $\sim 0.5$   $\mu\text{M}$ , whereas fluid-phase tES inhibited binding with an  $\text{IC}_{50}$  that was estimated by extrapolation to be at least 25  $\mu\text{M}$ . These results demonstrated that the affinity of PSGL-1 for tES was at least 50-fold lower than for tPS.





**FIG. 7. Interactions of tES and tPS with neutrophils or with  $^{125}\text{I}$ -PSGL-1.** A, neutrophils ( $2 \times 10^5/\text{well}$ ) were allowed to adhere for 30 min at  $22^\circ\text{C}$  to microtiter wells coated with the indicated amounts of tES or tPS. The plates were then sealed and centrifuged inverted for 5 min at  $200 \times g$ . Bound neutrophils were quantitated using a myeloperoxidase assay. B,  $^{125}\text{I}$ -PSGL-1 was incubated for 1 h in microtiter wells blocked with HSA or pre-coated with either tPS or tES ( $0.5 \mu\text{g}/\text{well}$ ) and then blocked with HSA. The samples were incubated in the presence or absence of 8 mM EDTA or 40  $\mu\text{g}/\text{ml}$  of the indicated mAb. The data represent the mean  $\pm$  S.D. of quadruplicate wells and are representative of three independent experiments.



**FIG. 8. Inhibition of  $^{125}\text{I}$ -PSGL-1 binding to immobilized tPS by fluid-phase tPS or tES.** A fixed concentration of  $^{125}\text{I}$ -PSGL-1 was incubated for 1 h in microtiter wells coated with tPS ( $50 \mu\text{l}$ ,  $10 \mu\text{g}/\text{ml}$ ) in the presence of increasing concentrations of either fluid-phase tPS or tES. The diamond represents basal adhesion of  $^{125}\text{I}$ -PSGL-1 to wells blocked with HSA but not coated with either selectin. The data represent the mean  $\pm$  S.D. of quadruplicate wells and are representative of two independent experiments.

#### DISCUSSION

We previously described a P-selectin glycoprotein ligand from human myeloid cells that is a disulfide-linked homodimer with few *N*-linked glycans but many clustered, sialylated *O*-linked oligosaccharides. The glycoprotein expresses the sLe<sup>x</sup> antigen, and removal of sialic acid eliminates recognition by P-selectin (6, 13). In this study, we have established that the polypeptide backbone of this ligand is identical to that of

PSGL-1, a glycoprotein ligand for P-selectin that was recently described by expression cloning from an HL-60 cell cDNA library (16). We have further characterized the structure and function of PSGL-1 from human neutrophils, using assays with the purified, radioiodinated glycoprotein. Our results suggest that this molecule presents an array of sialylated and fucosylated, *O*-linked poly-*N*-acetylactosamine that creates high affinity binding structures for P-selectin, but not for E-selectin.

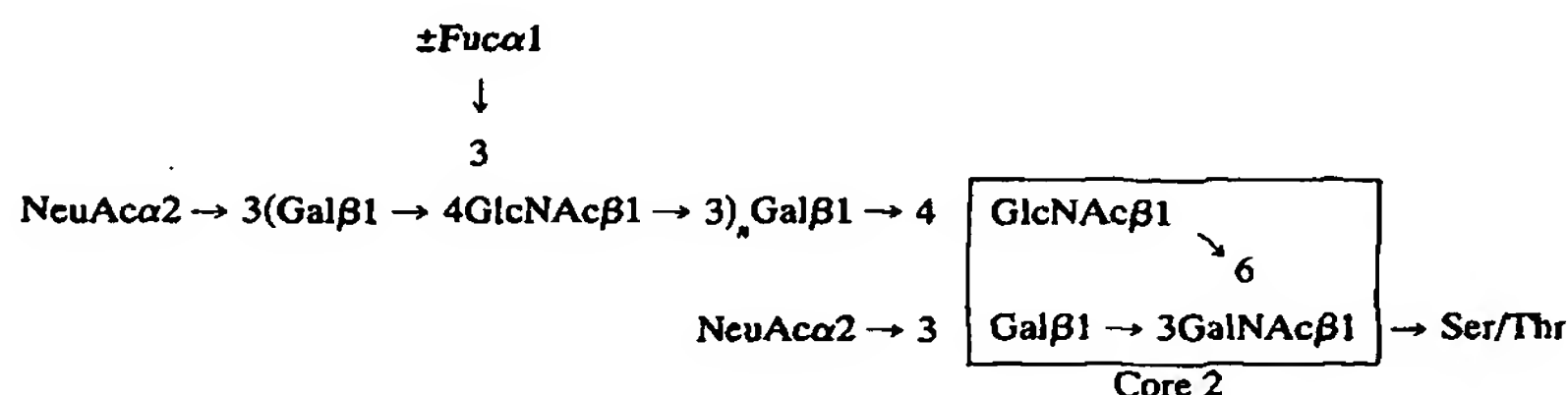
Although the great majority of the oligosaccharides on PSGL-1 are *O*-linked, there are a few *N*-linked glycans which we determined to be heterogeneous. All PSGL-1 molecules contained PNGaseF-sensitive *N*-linked oligosaccharides, but only 42% of the glycoprotein molecules bound to ConA, a lectin that binds well to high mannose *N*-linked glycans and less well to complex biantennary and hybrid *N*-linked chains. Based on the reductions in apparent molecular weight observed in SDS gels, PNGaseF removed at least two of the three possible *N*-linked glycans (16) from both the ConA-bound and -unbound fractions of PSGL-1. One chain on the ConA-unbound fraction of PSGL-1 may be resistant to PNGaseF, since the ConA-unbound material still migrated slightly slower in SDS gels than the ConA-bound material following treatment with PNGaseF. Alternatively, PNGaseF may have removed all the *N*-linked chains from both the ConA-bound and -unbound fractions, in which case the mobility difference may reflect heterogeneous *O*-linked glycosylation. The PNGaseF-sensitive *N*-linked glycans carried little or no Le<sup>x</sup> and sLe<sup>x</sup> and did not contribute measurably to the binding interaction with P-selectin. These results strongly suggest that the *O*-linked glycans display the relevant structures that P-selectin recognizes. It is highly improbable that a single PNGaseF-resistant *N*-linked glycan on a subpopulation of PSGL-1 is responsible for P-selectin recognition.

Because at least some of the *O*-linked oligosaccharides on PSGL-1 carry ( $\alpha$ 2-3)sialic acid and ( $\alpha$ 1-3)fucose, they must have relatively large, branched structures. In agreement with this prediction, we found that PSGL-1 carries few of the Ser/Thr-linked Gal $\beta$ 1-3GalNAc core 1 disaccharides that endo- $\alpha$ -*N*-acetylgalactosaminidase cleaves following removal of terminal sialic acids. These data are consistent with previous observations that less than 10% of the  $^3\text{H}$ -labeled *O*-linked oligosaccharides released by  $\beta$ -elimination from PSGL-1 were small as assessed by gel filtration (13).

The *O*-linked glycans on PSGL-1 include a heterogeneous group of poly-*N*-acetylactosamine. As measured by mobility on SDS gels and by binding to tomato lectin, endo- $\beta$ -galactosidase cleaved some of the poly-*N*-acetylactosamine at the internal  $\beta$ 1-4 linkage between the Gal and GlcNAc. Treatment with this enzyme also reduced the amount of Le<sup>x</sup> and sLe<sup>x</sup>, indicating that some of these structures were on poly-*N*-acetylactosamine. As measured by binding to tomato lectin, endo- $\beta$ -galactosidase cleaved other poly-*N*-acetylactosamine only after treatment with sialidase and fucosidase, demonstrating that they had sialic acid and fucose moieties at or near their terminus that prevented access to endo- $\beta$ -galactosidase. Because a significant portion of the PSGL-1 molecules still bound to tomato lectin after treatment with all three enzymes, there may be other *O*-linked poly-*N*-acetylactosamine chains that remained resistant to endo- $\beta$ -galactosidase because of branching or additional internal substitutions. Such structures might be responsible, at least in part, for the ability of PSGL-1 to bind to immobilized tPS after digestion with endo- $\beta$ -galactosidase.

Previous studies have shown that human myeloid cells express poly-*N*-acetylactosamine on some *O*-linked glycans. These structures occur almost exclusively on the branch attached to the C-6 of GalNAc as part of a core 2 *O*-linked oligosaccharide (49, 50). Of the  $\sim 80$  *O*-linked glycans on the major

sialomucin of human myeloid cells, leukosialin (CD43), only one or two contain poly-*N*-acetylactosamine, and approximately half of these have a terminal sLe<sup>x</sup> moiety (42), as shown in the following structure (Structure 1):



STRUCTURE 1

In contrast, PSGL-1 from neutrophils appears to contain a higher percentage of core 2, poly-*N*-acetylactosamine. Direct structural analysis will be required to determine the specific monosaccharides, linkages, and possible modifications such as sulfation, that constitute the O-linked glycans of PSGL-1. However, the current data suggest that some of these glycans may differ from those in leukosialin, implying that human myeloid cells can differentially glycosylate the polypeptide backbones of two distinct sialomucins. This differential glycosylation is functionally important, because P-selectin binds with high affinity to PSGL-1, but not to leukosialin (6).

The selectins require sialylation and fucosylation of most carbohydrate ligands for recognition (1-3). Similarly, PSGL-1 requires sialic acid (6) and fucose (16) to interact with P-selectin. Although PSGL-1 expresses the terminal sLe<sup>x</sup> tetrasaccharide, it is not clear whether it requires this structure *per se* for recognition. A simple model is that PSGL-1 presents many terminal sLe<sup>x</sup> moieties that increase the avidity with which it interacts with P-selectin. However, previous observations indicate that high concentrations of cell surface sLe<sup>x</sup> are not sufficient to confer optimal interactions with P-selectin (13, 44). It is more likely that additional modifications on each oligosaccharide chain are essential for optimal recognition. Specific features of several oligosaccharides in proximity may also present a clustered saccharide patch that promotes high affinity binding to P-selectin (13).

Sako *et al.* (16) generated a recombinant form of PSGL-1 that bound to P-selectin when expressed in COS cells co-transfected with an  $\alpha 1$ -3/4-fucosyltransferase, but not in COS cells lacking the fucosyltransferase. The relative molecular mass of the recombinant molecule was similar to that of PSGL-1 in human myeloid HL-60 cells, and the authors apparently concluded that the structure and function of recombinant PSGL-1 were equivalent to those of the native molecule. However, the specific oligosaccharide structures attached to native PSGL-1 and the glycosyltransferases required to construct these structures are unknown. Moreover, the  $\alpha 1$ -3/4-fucosyltransferase expressed in the COS cells to generate recombinant PSGL-1 differs from the fucosyltransferases in human myeloid cells (51, 52). Thus, it is not clear that the glycosylation, and hence the function, of recombinant PSGL-1 is identical to that of the native glycoprotein in human myeloid cells. Indeed, following sialidase digestion of recombinant PSGL-1, addition of endo- $\alpha$ -*N*-galactosaminidase caused a substantial increase in mobility of the glycoprotein (16), indicating that recombinant PSGL-1, in contrast to the native glycoprotein, contains many simple O-linked core 1 Gal $\beta$ 1-3GalNAc disaccharides that are susceptible to this enzyme.

Sako *et al.* (16) found that recombinant PSGL-1 no longer bound to P-selectin after treatment with sialidase and endo- $\alpha$ -

*N*-galactosaminidase, and concluded that O-linked oligosaccharides were important for recognition by P-selectin. Based on the data presented here, this conclusion appears to be correct.

However, the experiment of Sako *et al.* cannot be used to support a role for O-linked glycans in recognition. Prolonged digestion with sialidase is required to eliminate binding to native PSGL-1 (6), and the conditions for sialidase digestion of recombinant PSGL-1 used by Sako *et al.* reduced, but did not eliminate, binding of P-selectin. The loss of P-selectin binding after addition of endo- $\alpha$ -*N*-galactosaminidase, which releases only non-sialylated core 1 Gal $\beta$ 1-3GalNAc disaccharides, was more likely due to the previously observed contamination of this enzyme with sialidase, which would remove the remaining sialic acids required for recognition (6).

Sako *et al.* (16) also suggested that the *N*-linked oligosaccharides contributed to binding to P-selectin because recombinant P-selectin did not reprecipitate all of the PNGaseF-treated recombinant PSGL-1. However, these authors did not demonstrate whether the reprecipitation assay was quantitative by testing whether additional precipitation steps increased recovery of PSGL-1. The same group also concluded that *N*-linked glycosylation was important for P-selectin recognition because they observed that myeloid HL-60 cells treated with tunicamycin, an inhibitor of *N*-linked glycosylation, did not adhere to P-selectin-expressing cells (53). However, tunicamycin exerts indirect effects on cells, such as inhibiting the exit of some proteins from the endoplasmic reticulum, that make it difficult to interpret its effects on cell adhesion. We find no evidence that the *N*-linked glycans of PSGL-1 from human neutrophils play a role in recognition by P-selectin (this paper and Ref. 6).

E-selectin binds to small oligosaccharides containing sLe<sup>x</sup> as well as or better than P-selectin (54). Virtually all of the PSGL-1 molecules displayed sLe<sup>x</sup>, as measured by their ability to bind to a CSLEX-1 mAb affinity column. If terminal sLe<sup>x</sup> itself is the key recognition element, then E-selectin should bind to PSGL-1 as well as or even better than P-selectin. However, our results indicated that E-selectin bound to PSGL-1 with at least 50-fold lower affinity than did P-selectin. This finding demonstrates that E-selectin, like P-selectin, requires more than mere presentation of sLe<sup>x</sup> to bind with high affinity. It is possible that a small subset of the PSGL-1 molecules studied bound as well to E-selectin as to P-selectin, accounting for the overall weak binding. If so, the remainder of the PSGL-1 would have even less, if any, affinity for E-selectin.

Sako *et al.* (16) observed that recombinant PSGL-1 supported adhesion of transfected COS cells to cells expressing E-selectin as well as to cells expressing P-selectin, although the site densities of the adhesion molecules on the cells were not reported. Recently, Lenter *et al.* (55) reported that a molecule in murine neutrophils that appears to represent PSGL-1 was precipitated by both bivalent E-selectin and P-selectin Ig chimeras when coupled to protein A beads. Neither of these assays is sensitive to differences in the equilibrium affinities of E- and P-selectin



for PSGL-1. One interpretation of these studies is that PSGL-1 interacted with lower affinity with E-selectin than with P-selectin, but was still able to mediate static adhesion of PSGL-1-transfected COS cells to E-selectin-expressing cells or to bind to E-selectin immobilized on beads under the conditions used. Another interpretation of the studies of Sako *et al.* is that recombinant PSGL-1 is glycosylated significantly differently than the native glycoprotein, changing, and perhaps even reversing, the relative affinities of the molecule for E- and P-selectin. In either scenario, a difference in the affinities of PSGL-1 for E- and P-selectin might not be detectable in static cell adhesion experiments, particularly if the cells expressed high levels of PSGL-1 and the selectins. Whatever the affinity differences, it remains to be determined whether PSGL-1 on neutrophils forms physiologically relevant interactions with E- or P-selectin under the shear forces of the venular circulation.

In conclusion, our data indicate that PSGL-1 from human neutrophils displays complex, sialylated and fucosylated O-linked oligosaccharides, at least some of which appear to contain poly-N-acetylglucosamine with varying degrees of substitution. Presentation of these structures allows high affinity binding to P-selectin, but not to E-selectin. The sequences of these oligosaccharides, as well as their points of attachment to the polypeptide backbone, must be determined to understand the structural basis for the specific high affinity interaction of PSGL-1 with P-selectin. Comparison of the structure and function of recombinant PSGL-1 with that of the native glycoprotein in human myeloid cells will determine the extent of their similarities.

**Acknowledgments**—We thank Chris Titsworth and Sheryl Christoferson for excellent technical assistance, C. Wayne Smith for supplying monoclonal antibodies to E-selectin, Qun Zhou for measuring the binding of <sup>125</sup>I-PSGL-1 to the CSLEX-1 mAb column, Tom Zamborelli for PSGL-1 peptide synthesis, and Marynette Rihanek for titrating of the anti-peptide sera. We also thank Ajit Varki for critical reading of the manuscript.

## REFERENCES

- Lasky, L. A. (1992) *Science* **258**, 964–969
- Bevilacqua, M. P., and Nelson, R. M. (1993) *J. Clin. Invest.* **91**, 379–387
- McEver, R. P. (1994) *Curr. Opin. Immunol.* **6**, 75–84
- Foxall, C., Watson, S. R., Dowbenko, D., Fennie, C., Lasky, L. A., Kiso, M., Hasegawa, A., Asa, D., and Brandley, B. K. (1992) *J. Cell Biol.* **117**, 895–902
- Varki, A. (1992) *Curr. Opin. Cell Biol.* **257**, 257–266
- Moore, K. L., Stults, N. L., Diaz, S., Smith, D. L., Cummings, R. C., Varki, A., and McEver, R. P. (1992) *J. Cell Biol.* **118**, 445–456
- Lasky, L. A., Singer, M. S., Dowbenko, D., Imai, Y., Henzel, W. J., Grimley, C., Fennie, C., Gillett, N., Watson, S. R., and Rosen, S. D. (1992) *Cell* **69**, 927–938
- Levinovitz, A., Mühlhoff, J., Isenmann, S., and Vestweber, D. (1993) *J. Cell Biol.* **121**, 449–459
- Walcheck, B., Watts, G., and Jutila, M. A. (1993) *J. Exp. Med.* **178**, 853–863
- Baumhueter, S., Singer, M. S., Henzel, W., Hemmerich, S., Renz, M., Rosen, S. D., and Lasky, L. A. (1993) *Science* **262**, 436–438
- Berg, E. L., McEvoy, L. M., Berlin, C., Bargatze, R. F., and Butcher, E. C. (1993) *Nature* **366**, 695–698
- Norgard-Sumnicht, K. E., Varki, N. M., and Varki, A. (1993) *Science* **261**, 480–483
- Norgard, K. E., Moore, K. L., Diaz, S., Stults, N. L., Ushiyama, S., McEver, R. P., Cummings, R. D., and Varki, A. (1993) *J. Biol. Chem.* **268**, 12764–12774
- Ushiyama, S., Laue, T. M., Moore, K. L., Erickson, H. P., and McEver, R. P. (1993) *J. Biol. Chem.* **268**, 15229–15237
- Steininger, C. N., Eddy, C. A., Leimgruber, R. M., Mellors, A., and Welply, J. K. (1992) *Biochem. Biophys. Res. Commun.* **188**, 760–766
- Sako, D., Chang, X.-J., Barone, K. M., Vachino, G., White, H. M., Shaw, G., Veldman, G. M., Bean, K. M., Ahern, T. J., Furie, B., Cumming, D. A., and Larsen, G. R. (1993) *Cell* **75**, 1179–1186
- Merkle, R. K., and Cummings, R. D. (1987) *J. Biol. Chem.* **262**, 8179–8189
- McEver, R. P., and Martin, M. N. (1984) *J. Biol. Chem.* **259**, 9799–9804
- Geng, J.-G., Bevilacqua, M. P., Moore, K. L., McIntyre, T. M., Prescott, S. M., Kim, J. M., Bliss, G. A., Zimmerman, G. A., and McEver, R. P. (1990) *Nature* **343**, 757–760
- Mulligan, M. S., Varani, J., Kame, M. K., Lane, C. L., Smith, C. W., Anderson, D. C., and Ward, P. A. (1991) *J. Clin. Invest.* **88**, 1396–1406
- Shattil, S. J., Hoxie, J. A., Cunningham, M., and Brass, L. F. (1985) *J. Biol. Chem.* **260**, 11107–11114
- Laemmli, U. K. (1970) *Nature* **227**, 680–685
- Kozak, M. (1981) *Nucleic Acids Res.* **9**, 5233–5252
- Urlaub, G., and Chasin, L. A. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 4216–4220
- Bourdrel, L., Lin, C. H., Lauren, S. L., Elmore, R. H., Sugarman, B. J., Hu, S., and Wescott, K. R. (1993) *Protein Exp. Purif.* **4**, 130–140
- Turner, A. M., Zsebo, K. M., Martin, F., Jacobsen, F. W., Bennett, L. G., and Broudy, V. C. (1992) *Blood* **80**, 374–381
- Bevilacqua, M. P., Pober, J. S., Mendrick, D. L., Cotran, R. S., and Gimbrone, M. A., Jr. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 9238–9242
- Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem.* **182**, 319–326
- Laue, T. M. (1981) *Rapid Precision Interferometry for the Analytical Ultracentrifuge*. Ph. D. dissertation, University of Connecticut, Storres, CT
- Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. M. (1992) in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S., Rowe, A., and Horton, J. C., eds) pp. 90–125, Royal Society of Chemistry, London
- Yphantis, D. A. (1964) *Biochemistry* **3**, 297–317
- Johnson, M. L., and Frasier, S. G. (1985) *Methods Enzymol.* **117**, 301–342
- Stafford, W. (1992) *Anal. Biochem.* **203**, 295–301
- Zimmerman, G. A., McIntyre, T. M., and Prescott, S. M. (1985) *J. Clin. Invest.* **76**, 2235–2246
- Umemoto, J., Bhavanandan, V. P., and Davidson, E. A. (1977) *J. Biol. Chem.* **252**, 8609–8614
- Scudder, P., Uemura, K., Dolby, J., Fukuda, M. N., and Feizi, T. (1983) *Biochem. J.* **213**, 485–494
- Scudder, P., Hanfland, P., Uemura, K., and Feizi, T. (1984) *J. Biol. Chem.* **259**, 6586–6592
- Cummings, R. D. (1994) *Methods Enzymol.* **230**, 66–86
- Ogata, S., Muramatsu, T., and Kobata, A. (1975) *J. Biochem. (Tokyo)* **78**, 687–696
- Spooner, E., Fukuda, M., Klock, J. C., Oates, J. E., and Dell, A. (1984) *J. Biol. Chem.* **259**, 4792–4801
- Sano, M., Hayakawa, K., and Kato, I. (1992) *J. Biol. Chem.* **267**, 1522–1527
- Maemura, K., and Fukuda, M. (1992) *J. Biol. Chem.* **267**, 24379–24386
- Moore, K. L., Varki, A., and McEver, R. P. (1991) *J. Cell Biol.* **112**, 491–499
- Zhou, Q., Moore, K. L., Smith, D. F., Varki, A., McEver, R. P., and Cummings, R. D. (1991) *J. Cell Biol.* **115**, 557–564
- Phillips, M. L., Nudelman, E., Gaeta, F. C. A., Perez, M., Singhal, A. K., Hakomori, S., and Paulson, J. C. (1990) *Science* **250**, 1130–1132
- Walz, G., Aruffo, A., Kolanus, W., Bevilacqua, M., and Seed, B. (1990) *Science* **250**, 1132–1135
- Lowe, J. B., Stoolman, L. M., Nair, R. P., Larsen, R. D., Berhend, T. L., and Marks, R. M. (1990) *Cell* **63**, 475–484
- Polley, M. J., Phillips, M. L., Wayner, E., Nudelman, E., Singhal, A. K., Hakomori, S., and Paulson, J. C. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 6224–6228
- Fukuda, M., Carlsson, S. R., Klock, J. C., and Dell, A. (1986) *J. Biol. Chem.* **261**, 12796–12806
- Hanisch, F.-G., Uhlenbruck, G., Peter-Katalinic, J., Egge, H., Dabrowski, J., and Dabrowski, U. (1989) *J. Biol. Chem.* **264**, 872–883
- Goelz, S. E., Hession, C., Goff, D., Griffiths, B., Tizard, R., Newman, B., Chi-Rosso, G., and Lobb, R. (1990) *Cell* **63**, 1349–1356
- Sasaki, K., Kurata, K., Funayama, K., Nagata, M., Watanabe, E., Okta, S., Hanai, N., and Nishi, T. (1994) *J. Biol. Chem.* **269**, 14730–14737
- Larsen, G. R., Sako, D., Ahern, T. J., Shaffer, M., Erban, J., Sajer, S. A., Gibson, R. M., Wagner, D. D., Furie, B. C., and Furie, B. (1992) *J. Biol. Chem.* **267**, 11104–11110
- Nelson, R. M., Dolich, S., Aruffo, A., Cecconi, O., and Bevilacqua, M. P. (1993) *J. Clin. Invest.* **91**, 1157–1166
- Lenter, M., Levinovitz, A., Isenmann, S., and Vestweber, D. (1994) *J. Cell Biol.* **125**, 471–481

**A**bstract. The interaction of substance P with human blood T-lymphocytes, which stimulates T-lymphocyte proliferation, was quantified by both flow cytometric and direct binding assays. Fluorescence-detection flow cytometry recorded the binding of dichlorotriazinylamino-fluorescein-labeled substance P to  $21 \pm 10\%$  (mean  $\pm$  SD,  $n = 6$ ) and  $35 \pm 8\%$  ( $n = 2$ ) of human blood T-lymphocytes before and after stimulation with  $10 \mu\text{g/ml}$  of phytohemagglutinin, respectively. The suppressor-cytotoxic (leu 2a) and helper-inducer (leu 3a) subsets identified by phycoerythrin-labeled monoclonal antibodies contained substance P-reactive T-lymphocytes at respective mean frequencies of 10 and 18%. [ $^3\text{H}$ ]substance P bound rapidly and reversibly to a mean of  $7035 \pm 2850$  sites/T-lymphocyte, which exhibited a dissociation constant ( $K_D$ ) of  $1.85 \pm 0.70 \times 10^{-7}$  M (mean  $\pm$  SD,  $n = 5$ ). [D-Pro<sup>2</sup>,D-Phe<sup>7</sup>,D-Trp<sup>9</sup>]substance P inhibited the binding of dichlorotriazinylamino-fluorescein-labeled substance P and [ $^3\text{H}$ ]substance P to T-lymphocytes at concentrations that suppressed the proliferative response to substance P. Substance P (4–11), eledoisin, and substance K ( $\alpha$ -neurokinin), which all share with substance P the carboxy-terminal substituent -Gly-Leu-Met-NH<sub>2</sub>, were more potent than substance P (1–4) in inhibiting the binding of [ $^3\text{H}$ ]substance P to T-lymphocytes, suggesting the importance of this sequence in the interaction. Purified human blood B-lymphocytes, monocytes, polymorphonuclear leukocytes, and platelets, and cultured Hut 78 cutaneous lymphoma T-cells, Jurkat cells, Molt-4 lymphoblasts, and HL-60 and U-937 monocyte-like cells all showed only minimal specific binding of [ $^3\text{H}$ ]substance P. The recognition of substance

## Substance P Recognition by a Subset of Human T Lymphocytes

D. G. Payan, D. R. Brewster, A. Missirlian-Bastian, and E. J. Goetzl

Howard Hughes Medical Institute Laboratories and the Department of Medicine, University of California Medical Center, San Francisco, California 94143

P by T-lymphocytes provides one mechanism for selective modulation of immunity by sensory nerves.

### Introduction

Modulation of immunologic responses by elements of the central and peripheral nervous system appears to be attributable in part to the bidirectional effects of neuropeptides on distinct functions of lymphocytes. The results of in vitro studies of lymphocyte proliferation have shown that the responses to mitogens are suppressed by somatostatin (1, 2) and enhanced by substance P (SP)<sup>1</sup> (3) and  $\beta$ -endorphin (4). That different aspects of neuropeptide structure determine specifically the net effect on other lymphocyte functions was suggested by the capacity of  $\alpha$ -endorphin but not  $\beta$ -endorphin to inhibit T-lymphocyte-dependent antibody production (5), and of methionine-enkephalin and  $\beta$ -endorphin but not morphine or  $\alpha$ -endorphin to enhance the natural killer activity of lymphocytes (6). The interactions of neuropeptides with lymphocytes have been defined by direct analyses of the binding of methionine-enkephalin to T-lymphocytes (7) and of vasoactive intestinal polypeptide to both Molt-4 lymphoblasts and T-lymphocytes (8, 9).

SP is an undecapeptide of amino acid sequence Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH<sub>2</sub>, which has been identified in the central and peripheral nervous system and intestinal tract (10–13) and implicated in the mediation of hypersensitivity reactions by the detection of elevated levels of SP in sensory nerves supplying localized sites of chronic inflammation (14, 15). SP elicits or enhances functional responses of human mast cells, polymorphonuclear (PMN) leukocytes, T-lymphocytes, and guinea pig peritoneal macrophages at concentrations ranging from  $10^{-11}$  to  $10^{-5}$  M (3, 16–19). Furthermore, the stimulation of T-lymphocyte proliferation by SP was inhibited specifically by the otherwise immunologically inactive analogue [D-Pro<sup>2</sup>,D-Phe<sup>7</sup>,D-Trp<sup>9</sup>]SP (3, 20). The

Received for publication 11 November 1983 and in revised form 31 May 1984.

J. Clin. Invest.

© The American Society for Clinical Investigation, Inc.

0021-9738/84/10/1532/08 \$1.00

Volume 74, October 1984, 1532–1539

1. Abbreviations used in this paper: DTAF, dichlorotriazinylamino fluorescein; FACS, fluorescence-activated cell sorter; LTB<sub>4</sub>, leukotriene B<sub>4</sub>; M199-HPS and RPMI-HPS, medium 199 containing Hepes (25 mM, pH 7.4), penicillin (100 U/ml), and streptomycin (100  $\mu\text{g/ml}$ ); PE, phycoerythrin; PHA, phytohemagglutinin-M; SP, substance P; SP\*, substance P coupled to DTAF.



preparation and purification of fluorescein-labeled substance P (SP\*) now has permitted an evaluation by fluorescence-detection flow cytometry of the recognition of SP by a small subset of human blood T-lymphocytes, which manifest a specificity for SP\* and [<sup>3</sup>H]SP similar to that observed initially in studies of the effects of SP on T-lymphocyte function (3). Furthermore, the concurrent application of SP\* and of phycoerythrin-labeled monoclonal antibodies specific for antigenic determinants on functionally distinct subsets of T-lymphocytes now indicate that the SP\* reactive T-lymphocytes are distributed in both the suppressor-cytotoxic and helper-inducer subsets.

## Methods

Medium 199, RPMI-1640, sheep erythrocytes (Microbiological Associates Bioproducts, Walkersville, MD), 6 g% macromolecular dextran-70 in 0.15 M saline (Macrodex), Ficoll-Hypaque (Pharmacia Fine Chemicals, Inc., Piscataway, NJ), 4-(2-hydroxy-ethyl)-1-piperazine ethanesulfonic acid (Hepes), penicillin (1,000 U/ml), streptomycin (1,000 µg/ml; Gibco Laboratories, Grand Island, NY), phytohemagglutinin-M (PHA) (Difco Laboratories, Inc., Detroit, MI), *n*-butyl phthalate (Fisher Scientific Co., Pittsburgh, PA), dinonyl phthalate (ICN Pharmaceuticals, Inc., Plainview, NY), synthetic SP, SP (1-4), SP (4-11), eleodisin,  $\alpha$ -neurokinin (substance K), [D-Pro<sup>2</sup>,D-Phe<sup>7</sup>,D-Trp<sup>9</sup>]SP, somatostatin (1-14) (Peninsula Laboratories, Inc., Belmont, CA), [<sup>3</sup>H]SP (25-55 Ci/mmol) (New England Nuclear, Boston, MA), 1.5 ml conical polypropylene tubes (Sarstedt, Inc., Princeton, NJ), silica gel H of 250 µm thickness on 20 × 20-cm plates (Analtech, Inc., Newark, DE), phycoerythrin (PE)-conjugated monoclonal mouse antibodies to leu-3a and leu-2a (Becton-Dickinson, Inc., Mountainview, CA), dichlorotriazinylaminofluorescein (DTAF) (Research Organics, Inc., Cleveland OH), organic solvents which had been redistilled from glass (Burdick and Jackson Laboratories, Inc., Muskegon, MI), reagents for determining the amino acid composition of polypeptides (Beckman Instruments, Inc., Palo Alto, CA, and Pierce Chemical Co., Rockford, IL), human promyelocytic leukemia HL-60 and Jurkat cells (Dr. J. Stobo, University of California at San Francisco), cultures of Molt-4 lymphoblasts, U-937 monocyte-like cells, Hut 78 T-cells (American Type Culture Collection, Rockville, MD), and 1,25-dihydroxy vitamin D<sub>3</sub> [1,25(OH)<sub>2</sub>D<sub>3</sub>] (Hoffman-La Roche, Nutley, NJ) were obtained from the designated suppliers.

**Isolation of human T- and B-lymphocytes, PMN leukocytes, monocytes, and platelets.** Mixed leukocytes from heparin-anti-coagulated blood of normal subjects were centrifuged on Ficoll-Hypaque cushions to resolve mononuclear leukocytes from PMN leukocytes. The PMN leukocytes in the pellet were resuspended in 1 ml of AB-positive human serum that was diluted to 10 ml with distilled water at room temperature to lyse contaminating erythrocytes. After 20 sec, 40 ml of Medium 199 containing 25 mM Hepes (pH 7.4), 100 U/ml of penicillin, and 100 µg/ml of streptomycin (M199-HPS) was added, and the PMN leukocytes were recovered by centrifugation and washed twice with M199-HPS; the purity of the PMN leukocytes was >95% (21). The mononuclear leukocytes at the buffer and Ficoll-Hypaque interface were washed twice in M199-HPS and then were incubated with neuraminidase-treated fresh sheep erythrocytes to achieve rosetting of the T-lymphocytes (21). The mixtures were centrifuged on Ficoll-Hypaque cushions to separate the T-lymphocyte rosettes from less dense nonrosetting monocytes and B-lymphocytes. The erythrocytes were lysed by hypotonic exposure, and the T-lymphocytes were washed

and resuspended in M199-HPS. The purity of the T-lymphocytes was >95%, with <3% monocytes detected by a nonspecific esterase stain, and T-lymphocyte viability was always >97%, as determined by the exclusion of trypan blue dye (21).

The nonrosetting mixture of monocytes and B-lymphocytes was resuspended at a concentration of  $1 \times 10^6$ /ml of M199-HPS with 10% (vol:vol) human AB serum and incubated in 75 cm<sup>2</sup> plastic tissue culture flasks for 1 h at 37°C in 5% CO<sub>2</sub>:95% air to remove the adherent monocytes. The contents of the flasks were decanted; the nonadherent B-lymphocytes were washed and resuspended in M199-HPS and the incubation in plastic flasks repeated to remove residual adherent monocytes. The purity of the B-lymphocytes always was >90%, with <5% T-lymphocytes and 8% monocytes; B-lymphocyte viability was >94%, as determined by the exclusion of trypan blue dye (21). Monocytes were obtained by scraping the plastic 75 cm<sup>2</sup> flasks with a rubber spatula and washing the detached cells in M199-HPS. The purity of the monocytes was always >85% as assessed by a nonspecific esterase stain, and monocyte viability was always >90%.

Human platelets from normal subjects who had not taken aspirin or other platelet-active medication for at least 7 d were isolated from citrate-anticoagulated venous blood and washed three times on cushions of autologous erythrocytes, as described (22). The washed platelets were freed of erythrocytes by centrifugation at 10 g for 5 min and were resuspended in M199-HPS (pH 7.4) containing 0.1% (wt:vol) human serum albumin.

**Maintenance of Molt-4, U-937, Hut 78, Jurkat, and HL-60 cell cultures.** Cells were cultured at a density of  $\sim 5 \times 10^5$ /ml in 75 cm<sup>2</sup> plastic flasks in RPMI-1640 with L-glutamine, 15% (vol:vol) heat-inactivated fetal-calf serum, 25 mM Hepes (pH 7.4), penicillin (100 U/ml), and streptomycin (100 µg/ml) (RPMI-HPS) at 37°C in 5% CO<sub>2</sub>:95% air. The cultures were divided every 48-72 h and the cells used within 2 d of a subdivision. The cells were washed twice and resuspended in M199-HPS; viability was always >95%, as determined by the exclusion of trypan blue dye (21).

T-lymphocytes were cultured in RPMI-HPS for 48 h with and without 10 µg/ml of PHA and the degree of proliferation was quantified with a 0.2 ml portion of each suspension. The T-lymphocytes were diluted to  $1 \times 10^6$ /ml and transferred to wells of microtiter plates containing 1 µCi of [<sup>3</sup>H]thymidine (New England Nuclear) and the incubation continued for 8 h at 37°C in 5% CO<sub>2</sub>:95% air. The uptake of [<sup>3</sup>H]thymidine was quantified by trapping and washing the T-lymphocytes on glass fiber filters in a PHD cell harvester (Cambridge Technology, Inc., Cambridge, MA), and then counting the radioactivity as described (23). In two separate experiments, the uptake of [<sup>3</sup>H]thymidine by the unstimulated T-lymphocytes was  $253 \pm 67$  cpm (mean  $\pm$  SD) and by the PHA-stimulated T-lymphocytes was  $17,641 \pm 2345$  cpm.

**Induction of differentiation of U-937 and HL-60 cells by 1,25(OH)<sub>2</sub>D<sub>3</sub>.** U-937 and HL-60 cells were stimulated to differentiate into monocyte-like cells by 1,25(OH)<sub>2</sub>D<sub>3</sub> as described (24, 25) by incubation of suspensions of  $5 \times 10^5$ /ml of RPMI-HPS in 75 cm<sup>2</sup> tissue culture flasks with  $10^{-7}$  M 1,25(OH)<sub>2</sub>D<sub>3</sub> at 37°C in 5% CO<sub>2</sub>:95% air for 72 h. The contents of the flasks were decanted and the adherent differentiated cells removed with a rubber spatula were washed and resuspended in M199-HPS; viability of the adherent differentiated cells always exceeded 85%, as assessed by the exclusion of trypan blue. In two successive experiments, the extent of differentiation into monocytes was  $65 \pm 10\%$  (mean  $\pm$  SD) and  $80 \pm 5\%$  for the HL-60 and U-937 cells, respectively, as assessed by the increased percentage of adherent cells and the percentage which expressed nonspecific esterase activity (25).

**Preparation of DTAF-SP (SP\*).** To prepare fluorescent SP, 2.5–5.0 mg of DTAF and 0.5–1.0 mg of SP were reacted in 0.4–0.8 ml of 0.2 M sodium carbonate-buffered 0.15 M NaCl (pH 9.0) for 2 h at 37°C. The mixture was applied to one end of a 20 × 20-cm plate of 250 µm thick silica gel H that was developed in a sealed chamber with chloroform:methanol:glacial acetic acid (15:5:1, vol:vol:vol). Unreacted DTAF and SP, which were detected by ultraviolet light-induced fluorescence and ninhydrin staining, respectively, migrated as separate spots with  $R_F$  values of 0.91 and 0.05, respectively. SP coupled to DTAF (SP\*) migrated as a single spot with an  $R_F$  of 0.23. The SP\*-containing silica gel was scraped from the plate and eluted with four portions of 1 ml of methanol. The eluate then was dried with a continuous stream of  $N_2$  and resuspended in 0.4–0.8 ml of M199-HPS. To verify the amino acid composition of the SP\*, replicate portions were lyophilized in 9 × 150-mm glass test tubes, resuspended in 0.3 ml of constant boiling (5.7 M) HCl, and hydrolyzed as described (26). The amino acids were quantified with a Durrum D500 analyzer, utilizing norleucine as an internal standard (Dionex Co., Sunnyvale, CA) (26). For two different preparations, the relative amino acid composition of the SP\* was Lys 0.91 (1), Arg 0.99 (1), Gln 1.93 (2), Pro 2.11 (2), Gly 0.94 (1), Met 0.97 (1), Leu 1.02 (1), and Phe 2.2 (2), as compared with that expected for the sequence Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH<sub>2</sub> and indicated in the parentheses after each value.

**Assessment of the interaction of SP-DTAF (SP\*) with T-lymphocytes and other cells.** In each experiment,  $1 \times 10^7$  T-lymphocytes, monocytes, PMN leukocytes, or cultured cells in 0.2 ml M199-HPS were incubated for 40 min at 4°C with  $1-3 \times 10^{-7}$  M SP\* or unconjugated DTAF. The labeled T-lymphocytes were washed and resuspended in 1 ml of M199-HPS at 4°C immediately before introduction into a Becton-Dickinson fluorescence-activated cell sorter (FACS IV) equipped with a 2 W Argon laser (Spectra Physics, Mountainview, CA) operating at 400 mW that excited at 488 nm. The emitted light was passed through a long-pass filter and analyzed at 515 nm. Fluorescence intensity was measured on a scale of logarithmic amplitude which was calibrated so that a change of 57 channels was the equivalent of a 10-fold increase in fluorescence intensity. For each measurement of fluorescence,  $2 \times 10^5$  T-lymphocytes were counted. The interaction of SP\* with the T-lymphocytes was analyzed in terms of either two parameters (relative cell number, and relative fluorescence intensity) or three parameters (forward light scatter, relative cell number, and relative fluorescence intensity).

The interaction of SP with a specific subset of T-lymphocytes was quantified by dual color FACS analysis after sequential exposure of mixed T-lymphocytes to PE-labeled monoclonal leu-3a and leu-2a antibodies, which identify the helper-inducer and suppressor-cytotoxic subsets, respectively, and to SP coupled to DTAF (SP\*).  $1 \times 10^7$  T-lymphocytes in 0.2 ml of M199-HPS were incubated at 4°C for 45 min with 0.05 ml of either leu-3a or leu-2a bearing PE. The subset-labeled T-lymphocytes then were washed once and incubated with SP\* as described above. Dual color FACS analysis was carried out with a single 2 W Argon laser operating at 400 mW. Fluorescein staining was excited at 488 nm and the emitted light was passed through a long-pass filter and analyzed at 515 nm, while PE staining excited at 488 nm and was analyzed with a 500 nm beam splitter and two six-cavity band-pass filters (530/30 and 575/25) (Becton Dickinson, Inc.).

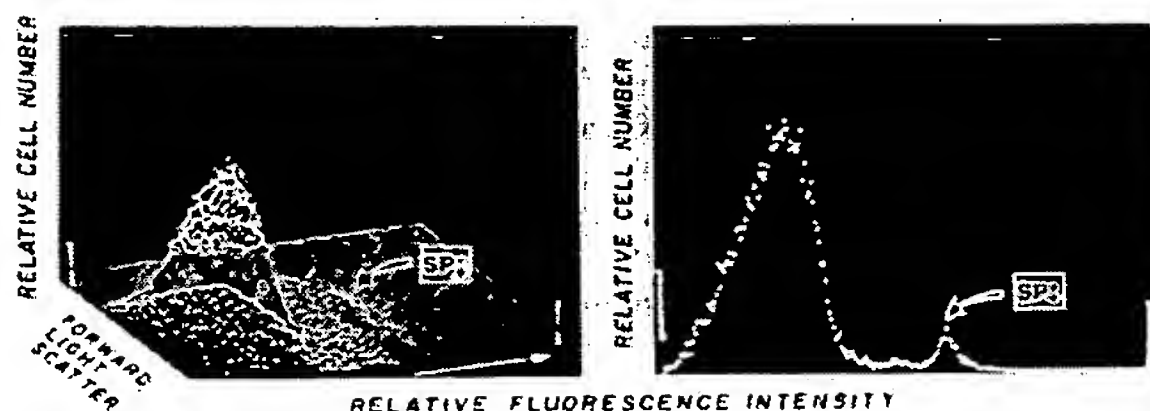
**Measurement of the binding of [<sup>3</sup>H]substance P to blood T-lymphocytes and other cells.** In each experiment, a concentration of [<sup>3</sup>H]SP ranging from 3 to 30 nM without and with different concentrations of SP, [D-Pro<sup>2</sup>,D-Phe<sup>7</sup>,D-Trp<sup>9</sup>]SP, SP substituent peptides,

substance K, eladoin, or somatostatin was incubated with duplicate suspensions of  $1 \times 10^7$  T-lymphocytes or other cells in a final volume of 0.3 ml for 40 min at 4°C. The amount of bound radioactivity was determined by sedimenting the T-lymphocytes in each suspension through a 0.3 ml layer of phthalate oils (27) in a 1.5 ml conical polypropylene tube that was centrifuged for 30 s at 8000 g in a Beckman microfuge B (Beckman Instruments, Inc.). The tip of the polypropylene tube containing the T-lymphocyte pellet was cut off with a razor blade, and the contents of the tip were resuspended with a pasteur pipette in 3 ml of Hydrofluor for the quantification of radioactivity in the pellet.

The total number of moles of SP bound to the T-lymphocytes was determined by dividing the cpm bound to the cells by the specific activity of [<sup>3</sup>H]SP. The amount of radioactivity bound in the presence of  $3.0 \times 10^{-8}$  to  $1.0 \times 10^{-4}$  M nonradioactive SP was divided by the same value for the specific activity of [<sup>3</sup>H]SP to determine the level of nonspecific binding of SP. The number of moles of SP specifically bound to the T-lymphocytes was calculated by subtracting the nonspecific binding from the total binding.

## Results

**Characteristics of the interaction of SP-DTAF (SP\*) with human blood T-lymphocytes.** T-lymphocytes were labeled with SP\* and the distribution of labeling analyzed with a FACS to assess whether SP is recognized by all of the T-lymphocytes or only a limited subpopulation. The flow cytometric analyses demonstrated a small peak of positively fluorescent cells (Fig. 1), which represents the T-lymphocytes that bind SP\*. In six consecutive experiments with T-lymphocytes purified from different normal subjects, the number of cells labeled with SP\* accounted for  $21 \pm 10\%$  (mean  $\pm$  SD) of the total T-lymphocytes. In contrast, DTAF alone failed to reveal a specifically staining population of cells. The specificity of labeling of T-lymphocytes with SP\* was examined using a similar experimental design, but including replicate suspensions incubated with a mixture of SP\* and  $10^{-5}$  M [D-Pro<sup>2</sup>,D-Phe<sup>7</sup>,D-Trp<sup>9</sup>]SP, which prevents the enhancement of T-lymphocyte function by SP (3). In the



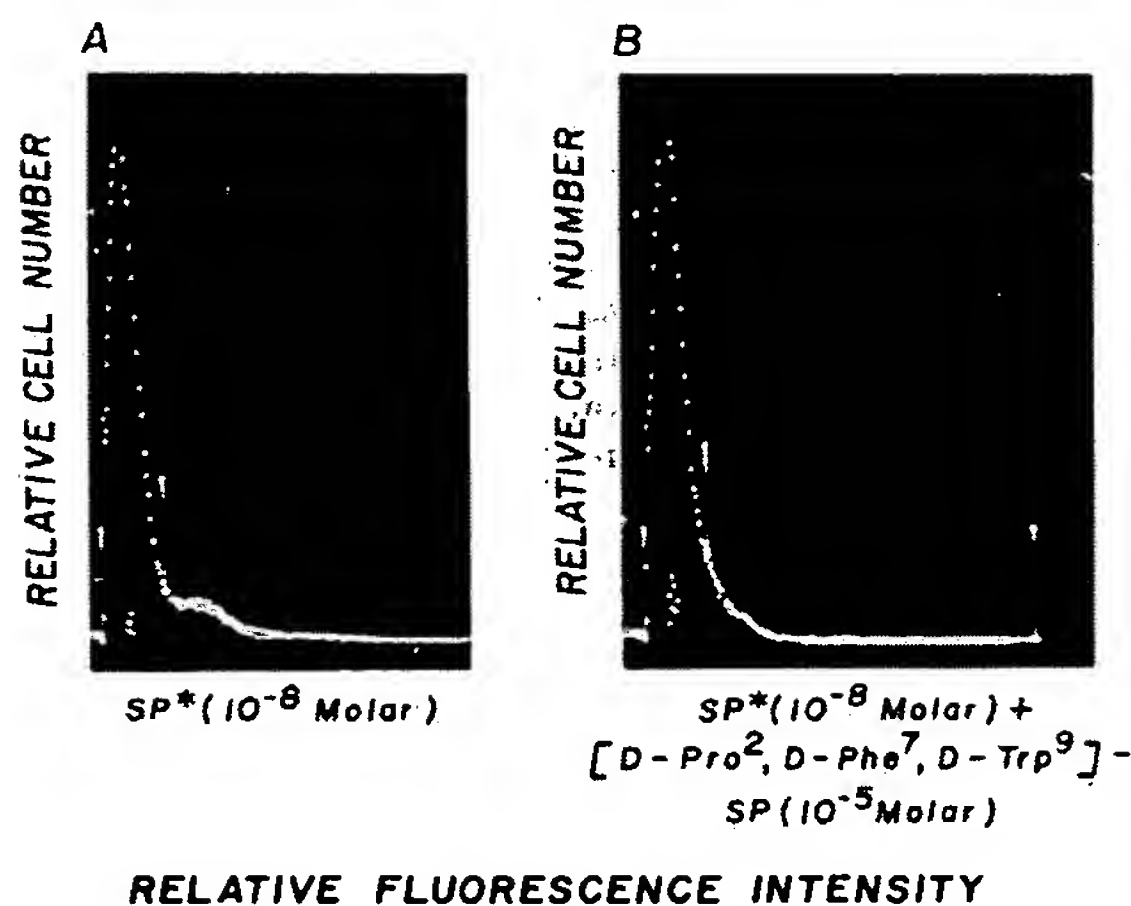
**Figure 1.** Flow cytometric analysis of the binding of SP\* to human T-lymphocytes.  $1 \times 10^7$  T-lymphocytes in 0.2 ml of M199-HPS were incubated for 40 min at 4°C with  $1-3 \times 10^{-7}$  M SP\*, washed, and analyzed in a FACS IV. Fluorescence intensity is depicted on a log amplitude scale. The subset of T-lymphocytes reactive with SP is designated SP\*.



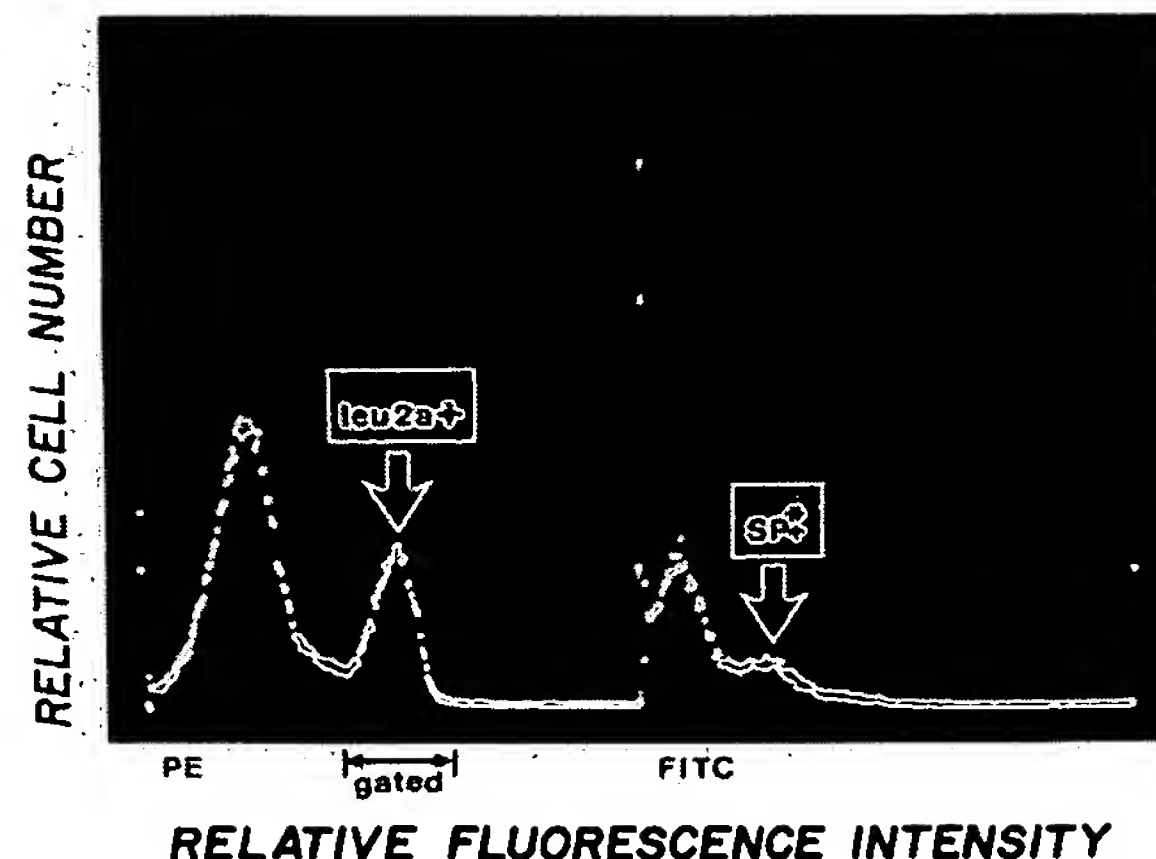
absence of [D-Pro<sup>2</sup>,D-Phe<sup>7</sup>,D-Trp<sup>9</sup>]SP, a positively fluorescent peak of T-lymphocytes labeled with SP\* is detected, representing 15% of the total T-lymphocytes (Fig. 2 A). The peptide antagonist of the effects of SP on T-lymphocytes eliminated completely any labeling of T-lymphocytes with SP\* (Fig. 2 B).

The binding of SP\* to T-lymphocytes, which had been stimulated by PHA to proliferate, was evaluated in order to examine the responsiveness of the SP-reactive subset and the effects of expansion of the subset on the specificity of binding of SP. At the end of the period of exposure to PHA, T-lymphocyte viability was 85±5% (mean±SD, *n* = 2) and 81±7% for the unstimulated and PHA-stimulated T-lymphocytes, respectively, as assessed by the exclusion of trypan blue. The labeling of PHA-stimulated T-lymphocytes with SP\* increased to 35±8% (mean±SD, *n* = 2), as compared with 23±6% for unstimulated T-lymphocytes from the same donors. The binding of SP\* by PHA-stimulated T-lymphocytes was reversed completely by 10<sup>-5</sup> M [D-Pro<sup>2</sup>,D-Phe<sup>7</sup>,D-Trp<sup>9</sup>]SP, as with unstimulated T-lymphocytes. In two consecutive experiments with SP\*, PMN leukocytes, monocytes, Molt-4 lymphoblasts, Hut 78 cells, and Jurkat cells failed to develop specific fluorescence, indicating a lack of recognition of SP.

Dual immunofluorescent staining of T-lymphocytes with SP\* and subset-specific monoclonal antibodies conjugated with PE demonstrated SP\*-reactive T-lymphocytes in both subsets (Figs. 3 and 4). In the suppressor-cytotoxic subset identified



**Figure 2.** Flow cytometric analysis of the specificity of binding of SP\* to human T-lymphocytes.  $1 \times 10^7$  T-lymphocytes were incubated with SP\* as described in Fig. 1. In (A), the binding of  $1 \times 10^{-8}$  M SP\* was subjected to two parameter analysis; in (B), T-lymphocytes were incubated with both  $1 \times 10^{-8}$  M SP\* and  $10^{-5}$  M [D-Pro<sup>2</sup>,D-Phe<sup>7</sup>,D-Trp<sup>9</sup>]SP and an identical analysis performed. With the addition of [D-Pro<sup>2</sup>,D-Phe<sup>7</sup>,D-Trp<sup>9</sup>]SP, the small positively staining peak in A is no longer present in B.

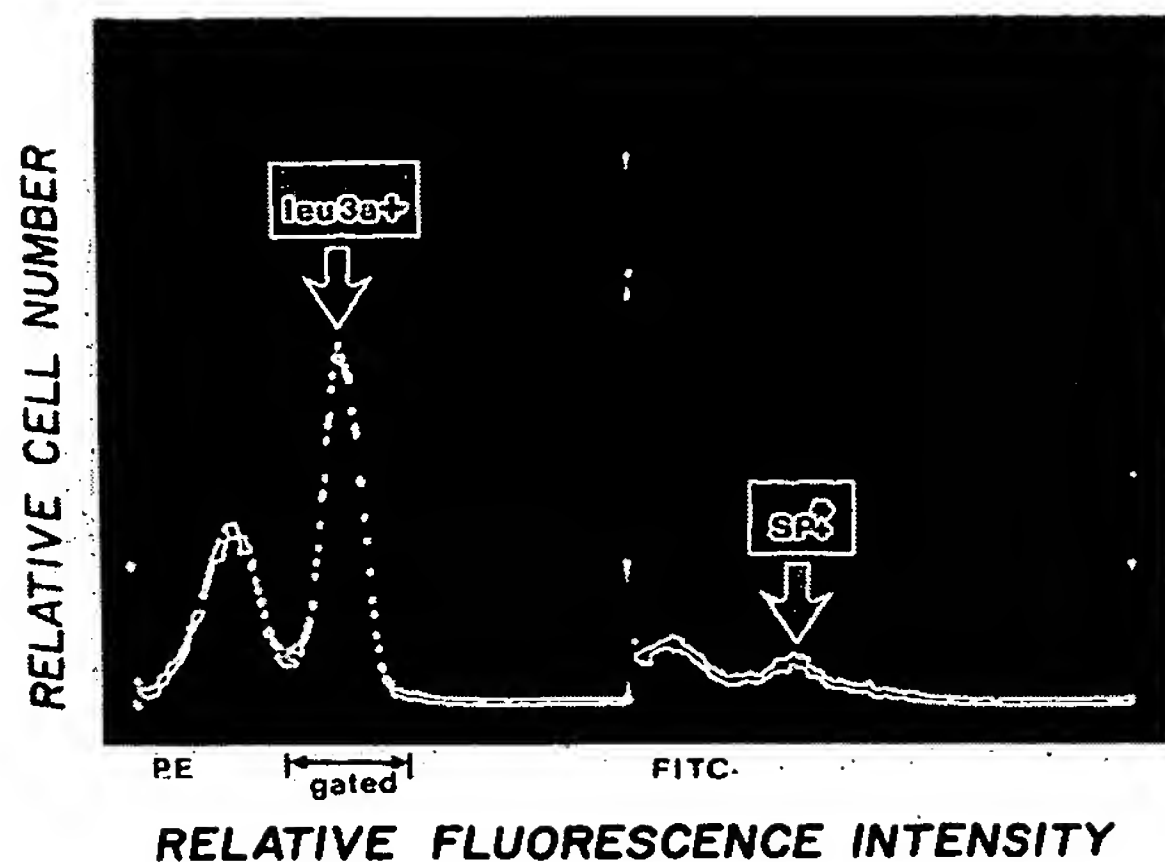


**Figure 3.** Two-color flow cytometric analysis of leu 2a-PE (suppressor-cytotoxic) expression and binding of SP-DTAF (SP\*) to human blood T-lymphocytes. T-lymphocytes were stained sequentially with anti-leu 2a-PE and SP\* and analyzed with a FACS IV with two parameters. The positively-staining anti-leu 2a-PE T-lymphocytes (left) were analyzed selectively (gated peak) and examined for the presence of SP\* reactivity. Within the leu 2a-positive peak, there are both nonreactive and SP\*-reactive (9.5%) (right) T-lymphocytes. The positively-staining anti-leu 2a-PE T-lymphocytes constituted 26±5% (mean±SD, *n* = 3) of the total.

by leu 2a-PE, 9.5±3.0% (mean±range, *n* = 3) of the T-lymphocytes bound SP\* (Fig. 3), while in the helper-inducer subset identified by leu 3a-PE, 18.0±5.0% (mean±range, *n* = 3) of the T-lymphocytes bound SP\* (Fig. 4).

**Characteristics of the binding of [<sup>3</sup>H]SP to human blood T-lymphocytes.** The time-course of association of [<sup>3</sup>H]SP with purified T-lymphocytes was assessed by incubating duplicate suspensions of  $1 \times 10^7$  T-lymphocytes with 10 nM [<sup>3</sup>H]SP for 0.5–40 min at 4°C. The specific binding of [<sup>3</sup>H]SP by T-lymphocytes increased with time and reached a plateau after ~9–12 min (Fig. 5). The time-course of dissociation of specifically bound [<sup>3</sup>H]SP from T-lymphocytes was examined by incubating replicate suspensions of T-lymphocytes for 40 min at 4°C with 10 nM [<sup>3</sup>H]SP to achieve saturation, washing with cold buffer to remove fluid phase [<sup>3</sup>H]SP, and resuspending in buffer with 10<sup>-4</sup> M (10,000-fold excess) nonradioactive SP. Specifically bound [<sup>3</sup>H]SP diminished rapidly with time, so that ~50 and 80% were dissociated after 5 and 30 min, respectively (Fig. 5).

A representative Scatchard plot (Fig. 6) of the concentration-dependence of [<sup>3</sup>H]SP binding to T-lymphocytes reveals a linear relationship with a single dissociation constant (*K<sub>D</sub>*) of  $0.83 \times 10^{-7}$  M and a mean of 4820 specific binding sites for SP/T-lymphocyte. The data from this and four other identical analyses of the binding of [<sup>3</sup>H]SP to T-lymphocytes from different normal subjects exhibit some donor-dependent vari-



**Figure 4.** Two-color flow cytometric analysis of leu 3a-PE (helper-inducer) expression and binding of SP-DTAF (SP\*) to peripheral blood T-lymphocytes. T-lymphocytes were stained with anti-leu 3a-PE and SP\* and analyzed with a FACS IV as in Fig. 3. The positively staining anti-leu 3a-PE T-lymphocytes (left) were analyzed selectively (gated peak) and examined for the presence of SP\* reactivity. Within the leu 3a-positive peak, there were both nonreactive and SP\* reactive (18%) (right) T-lymphocytes. The positively staining anti-leu 3a-PE T-lymphocytes constituted  $60 \pm 8\%$  (mean  $\pm$  SD,  $n = 3$ ) of the total.

ation (Table I). The mean  $K_D \pm$  SD was  $1.85 \pm 0.70 \times 10^{-7}$  M and the mean density of binding sites for SP  $\pm$  SD was  $7035 \pm 2850$ /T-lymphocyte. The level of nonspecific binding of [ $^3$ H]SP after a 40-min incubation at  $4^\circ\text{C}$  for five separate experiments was  $35 \pm 15\%$  (mean  $\pm$  SD) of the total binding.

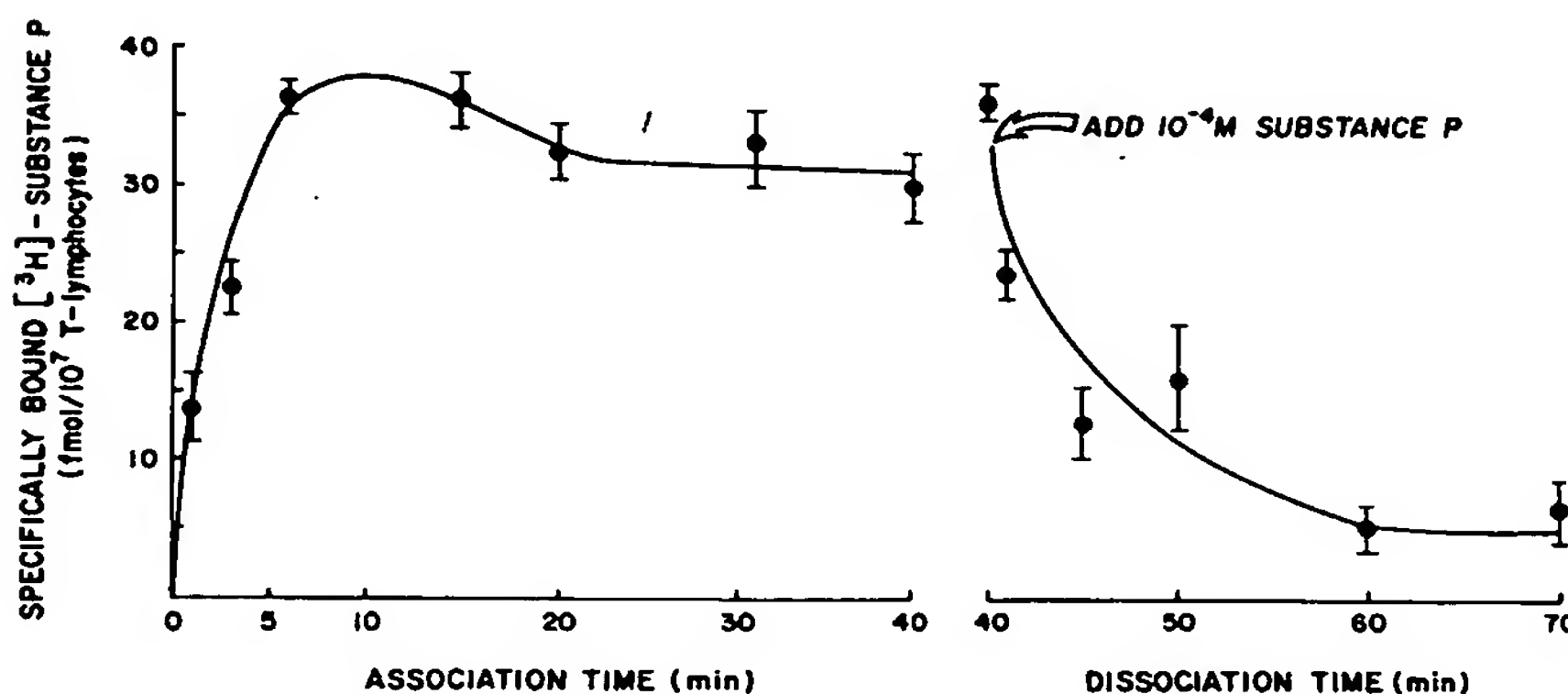
**Structural determinants of the binding of SP to human T-lymphocytes.** The peptide requirements for the binding of [ $^3$ H]SP to T-lymphocytes was evaluated by examining the effects on binding of structurally distinct neuropeptides and

substituents of SP (Fig. 5). The SP antagonist (19) [D-Pro<sup>2</sup>, D-Phe<sup>7</sup>, D-Trp<sup>9</sup>]SP competitively inhibited the binding of [ $^3$ H]SP to T-lymphocytes (Fig. 7) at concentrations known to suppress the stimulatory effects of SP on T-lymphocyte function (3). SP (4–11), which stimulates [ $^3$ H]thymidine incorporation by T-lymphocytes (3), also competitively inhibited the binding of [ $^3$ H]SP to T-lymphocytes. Eledoisin, which has the same amino acids as SP at positions 2, 7, and 9–11, and substance K ( $\alpha$ -neurokinin) (28), which has the same amino acids as SP at positions 3, 7, and 9–11, exhibited  $\sim 10$ –20% of the potency of SP assessed at 20% displacement of binding of [ $^3$ H]SP and 10% and 0.1%, respectively, of the potency of SP at 40% displacement (Fig. 7). In contrast, similar concentrations of the tetrapeptide SP (1–4) and of somatostatin failed to inhibit significantly the binding of [ $^3$ H]SP; the concentration of somatostatin required for 50% inhibition of specific binding of [ $^3$ H]SP was 1,000-fold greater than for SP.

**Cellular specificity of the binding of SP.** The binding of [ $^3$ H]SP by purified human blood B-lymphocytes, monocytes, PMN leukocytes, and platelets, and by cultured Jurkat, Hut 78, Molt-4, and differentiated U-937 and HL-60 cells was examined in two experiments (Table II). None of the non-T-lymphocytes demonstrated specific binding of [ $^3$ H]SP which exceeded 21% of the total binding. In contrast, the total binding of [ $^3$ H]SP by PHA-stimulated and unstimulated blood T-lymphocytes, which achieved a level similar to that of the cultured cells, consisted of over 60% specific binding as assessed by the addition of a 10,000-fold molar excess of unlabeled SP (Table II).

## Discussion

The capacity of neuropeptides and other mediators to modify immune function specifically is dependent on recognition of the mediators by T-lymphocytes which express immunoregulatory activity. The application of two distinct techniques employing SP labeled by different methods demonstrated the



**Figure 5.** Time-course of specific binding of [ $^3$ H]SP to human T-lymphocytes. [ $^3$ H]SP was incubated with  $1 \times 10^7$  T-lymphocytes for different periods of time and the levels of specific binding were quantified. Each point and bracket represent the mean  $\pm$  SD of three experiments done in duplicate.

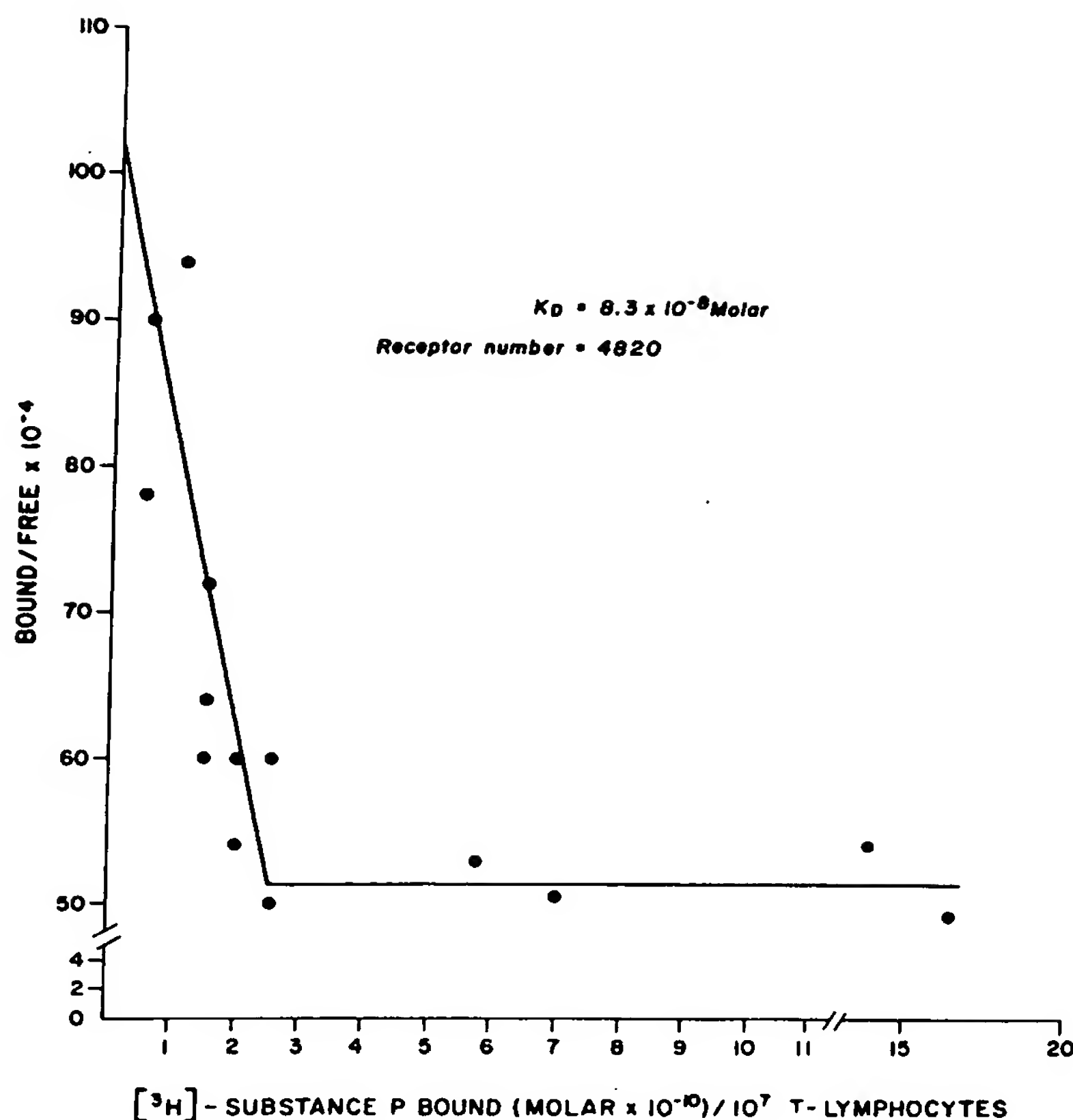


Figure 6. Scatchard plot of the concentration-dependence of  $[^3\text{H}]\text{SP}$  binding to human T-lymphocytes. The data are based on total  $[^3\text{H}]\text{SP}$  bound by  $1 \times 10^7$  T-lymphocytes in the presence of increasing quantities of nonradioactive SP. Each point represents the mean of duplicate measurements for experiment number 1 (see Table I). The horizontal portion of the curve represents the amount of nonspecific binding of  $[^3\text{H}]\text{SP}$  to T-lymphocytes.

specificity of the interaction of SP with human T-lymphocytes. A fluorescent derivative of SP, designated SP\*, labeled  $21 \pm 10\%$  (mean  $\pm$  SD,  $n = 6$ ) of human T-lymphocytes as assessed by flow cytometry (Fig. 1). The binding of  $[^3\text{H}]\text{SP}$  by similarly

Table I. Binding of  $[^3\text{H}]\text{SP}$  to Human Blood T-Lymphocytes

Experiment*	Specific binding sites/ T-lymphocyte†	$K_D$ (Molar $\times 10^{-7}$ )‡
1	4820	0.83
2	7230	3.33
3	9033	1.40
4	10480	2.22
5	3614	1.51
Mean $\pm$ SD	7035 $\pm$ 2850	1.85 $\pm$ 0.70 $\times 10^{-7}$ M

\* Experiments were done on blood T-lymphocytes from five different normal donors.

† The  $K_D$  and number of binding sites were determined by Scatchard analysis as described in the text. Each experiment was done in duplicate.

purified human T-lymphocytes revealed a single class of receptors with a  $K_D$  of  $1.85 \pm 0.70 \times 10^{-7}$  M (mean  $\pm$  SD,  $n = 5$ ) (Fig. 6, Table I). The binding of  $[^3\text{H}]\text{SP}$  to T-lymphocyte receptors achieved equilibrium rapidly and was reversed rapidly by the addition of excess nonradioactive SP (Fig. 5). The specificity of the interaction of SP with T-lymphocytes was confirmed by complete inhibition of both SP\* fluorescent staining and  $[^3\text{H}]\text{SP}$  binding by  $[\text{D-Pro}^2, \text{D-Phe}^7, \text{D-Trp}^9]\text{SP}$  (Fig. 7) at concentrations which inhibit the effects of SP on T-lymphocyte function (3). The competitive inhibition of  $[^3\text{H}]\text{SP}$  binding by other peptides revealed that SP (4-11), which stimulates T-lymphocyte proliferation with a concentration dependence similar to SP (3), displaced  $[^3\text{H}]\text{SP}$  at equimolar concentrations (Fig. 7). Two peptides of the tachykinin family, eledoisin and substance K, which share the same carboxy-terminal tripeptide and two other amino acids with SP, also inhibit the binding of  $[^3\text{H}]\text{SP}$  to T-lymphocytes at higher concentrations (Fig. 7). The structural requirements for binding of SP to T-lymphocytes suggest that the carboxy-terminal sequence -Gly-Leu-Met-NH<sub>2</sub>, which contributes to the stabilization of the structure of SP (29) and Phe<sup>7</sup> but not the

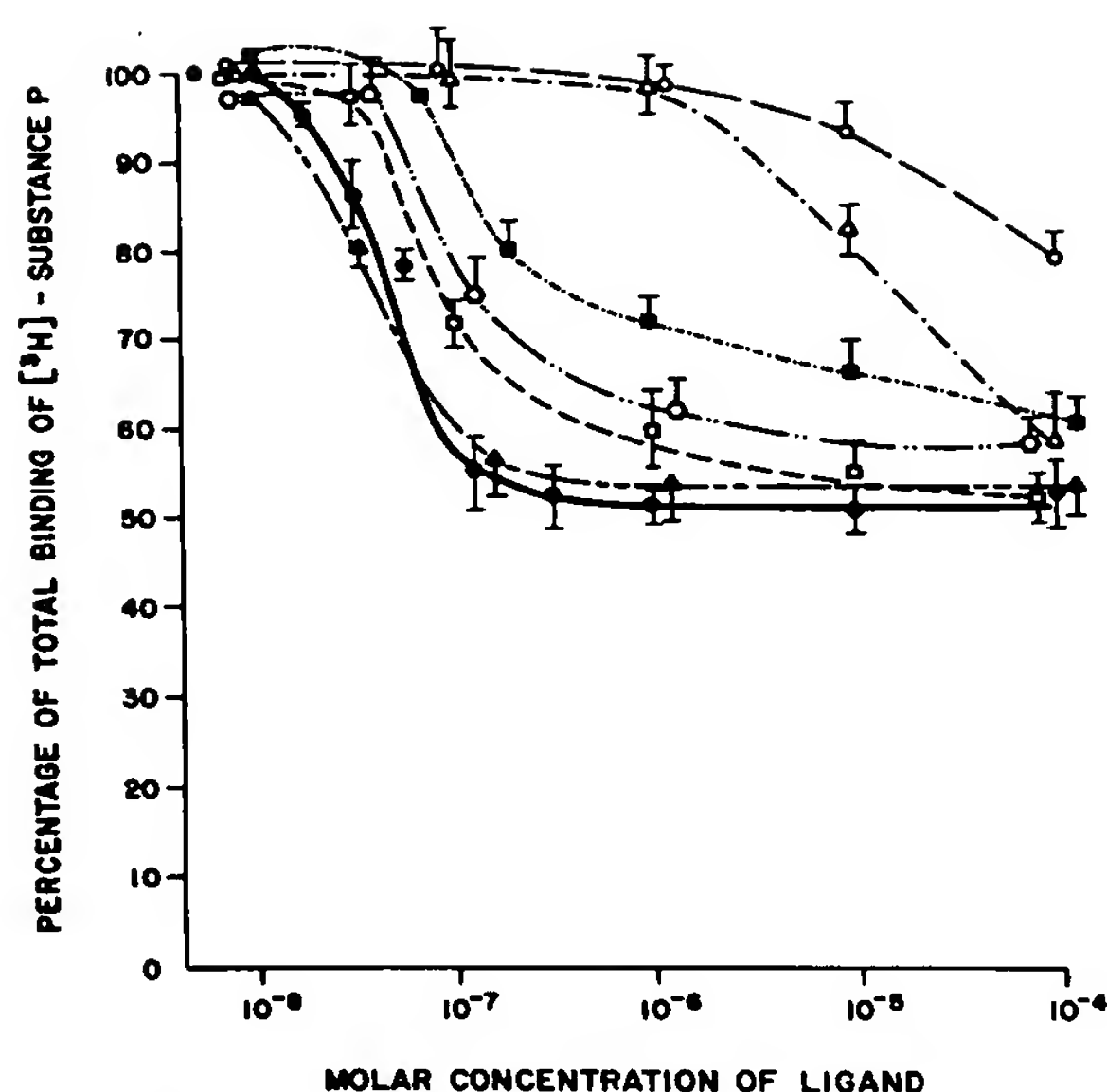


Figure 7. Structural determinants of binding of  $[^3\text{H}]\text{SP}$  to human T-lymphocytes. Each point and bracket depicts the mean  $\pm$  SD of the percentage of total binding of  $[^3\text{H}]\text{SP}$  inhibited specifically by increasing concentrations of the nonradioactive peptides listed below, and  $n$  equals the number of experiments.  $\bullet$ ,  $n = 4$ , SP Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH<sub>2</sub>;  $\circ$ ,  $n = 2$ , SP (1-4) Arg-Pro-Lys-Pro;  $\blacktriangle$ ,  $n = 2$ , SP (4-11) Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH<sub>2</sub>;  $\triangle$ ,  $n = 2$ , elodisin pGlu-Pro-Ser-Lys-Asp-Ala-Phe-Ile-Gly-Leu-Met-NH<sub>2</sub>;  $\blacksquare$ ,  $n = 2$ , substance K His-Lys-Thr-Asp-Ser-Phe-Val-Gly-Leu-Met-NH<sub>2</sub>;  $\triangle$ ,  $n = 2$ , somatostatin (1-14) Ala-Gly-Cys-Lys-Asn-Phe-Phe-Trp-Lys-Thr-Phe-Thr-Ser-Cys;  $\square$ ,  $n = 2$ , [D-Pro<sup>2</sup>,D-Phe<sup>7</sup>,D-Trp<sup>9</sup>]SP.

amino-terminal tetrapeptide, determine the affinity and specificity of the interaction.

Concurrent identification of helper-inducer and suppressor-cytotoxic subsets of T-lymphocytes with PE-conjugated monoclonal antibodies to different leu antigens of the respective subsets permitted assignment of the SP-DTAF (SP\*) fluorescence to T-lymphocytes in both subsets (Fig. 3 and 4). SP\* was recognized by a mean of 18 and 9.5% of the helper-inducer and suppressor-cytotoxic subsets of T-lymphocytes, respectively. The cellular specificity of the binding of SP was established by demonstrating a greater degree of specific binding of  $[^3\text{H}]\text{SP}$  to T-lymphocytes than to B-lymphocytes, monocytes, PMN leukocytes, platelets, and differentiated monocytes of the U-937 and HL-60 lines (Table II). SP, at concentrations of  $10^{-11}$ – $10^{-6}$  M had no effect on the uptake of  $[^3\text{H}]\text{thymidine}$  by the lymphoid cells Molt-4, Hut-78, and Jurkat in culture, and  $[^3\text{H}]\text{SP}$  failed to bind the cultured lymphoid cells.

The demonstration of labeling of only a mean of 21% of mixed T-lymphocytes with SP\* permitted an estimation of the

Table II. Binding of  $[^3\text{H}]\text{SP}$  to Different Types of Cells

Cell type	Total binding ( $10^{-8}$ M $[^3\text{H}]\text{SP}$ )	Specific binding†
T-lymphocytes*	1273 $\pm$ 126	766 $\pm$ 96
T-lymphocytes stimulated with PHA*	1626 $\pm$ 80	994 $\pm$ 87
B-lymphocytes*	966 $\pm$ 107	19 $\pm$ 65
Monocytes‡	1451 $\pm$ 161	97 $\pm$ 130
PMN leukocytes‡	972 $\pm$ 60	158 $\pm$ 67
Platelets‡	2817 $\pm$ 90	119 $\pm$ 120
Molt-4 lymphoblasts	3810 $\pm$ 368	369 $\pm$ 248
Jurkat cells	2641 $\pm$ 210	174 $\pm$ 190
Hut 78 T-cells	2185 $\pm$ 162	189 $\pm$ 141
U-937 monocytes, differentiated§	846 $\pm$ 35	196 $\pm$ 86
HL-60 cells, differentiated§	724 $\pm$ 117	178 $\pm$ 54

\* Experiments utilized T- and B-lymphocytes isolated from five separate normal donors. T-lymphocytes were stimulated with 10  $\mu\text{g}/\text{ml}$  of PHA for 48 h.

‡ Experiments utilized PMN leukocytes, monocytes, and platelets from two separate normal donors.

§ Two separate experiments were performed with cells stimulated to differentiate by  $1,25(\text{OH})_2\text{D}_3$ .

|| Mean cpm bound  $\pm$  SD/ $1 \times 10^7$  cells for duplicate measurements with cells from each donor, except in the case of platelets where  $1 \times 10^9$  cells in duplicate were used.

† Specific binding = total binding ( $10^{-8}$  M  $[^3\text{H}]\text{SP}$ ) minus nonspecific binding ( $10^{-8}$  M  $[^3\text{H}]\text{SP} + 10^{-4}$  M SP).

number of SP receptors. The Scatchard plot of the results of binding of  $[^3\text{H}]\text{SP}$  to T-lymphocytes revealed a mean of 7035 SP receptors/T-lymphocyte, if all cells were assumed to bind SP (Fig. 6, Table I). Utilizing the flow cytometric data to define a subset of SP-reactive T-lymphocytes which amount to 21% of the total population led to an estimate of  $\sim 35,000$  receptors/T-lymphocyte in the subset which recognized SP. The relevance of the binding data to the effects of SP on T-lymphocytes also was suggested by the similarity of the  $K_D$  to the concentration of  $0.5 \times 10^{-7}$  M SP required to achieve 50% of the maximal stimulation of T-lymphocyte proliferation (3).

The selective interaction of SP with a quantitatively small subset of T-lymphocytes is reminiscent of the binding of C3a peptides (21) and leukotriene B<sub>4</sub> (LTB<sub>4</sub>) (30) to only  $\sim 40$  and 11%, respectively, of human blood T-lymphocytes. The distribution of SP\*-reactive T-lymphocytes between the helper-inducer (leu 3a+) and suppressor-cytotoxic (leu 2a+) subsets resembles more that of LTB<sub>4</sub>-reactive lymphocytes within the same general subsets (30). Although LTB<sub>4</sub> reacted with both subsets, a greater percentage of LTB<sub>4</sub>-reactive lymphocytes were suppressor-cytotoxic cells (14%) than helper-inducer cells (8%), with the net effect of LTB<sub>4</sub> on proliferation being inhibition (23). In contrast, a greater percentage of SP-reactive



T-lymphocytes were helper-inducer cells (18%) compared with suppressor-cytotoxic cells (9%), and the net effect of SP on proliferation is stimulation (3). However, the presence in the helper-inducer subset of suppressor-inducer T-lymphocytes emphasizes the difficulty of meaningfully explaining a functional outcome from the results of analyses with a limited number of monoclonal antibody reagents. The recognition of C3a peptides and somatostatin by a minority of T-lymphocytes also has substantial functional consequences, as manifested by significant suppression of proliferation (1, 21) and other activities (21). The in vitro findings suggest that the net effect on immunological responses of neuropeptides released from peripheral nerves may depend, in part, on the balance of the stimulatory effects of SP and the inhibitory effects of somatostatin.

### Acknowledgments

This work was supported in part by National Institutes of Health grants 1P01 AI19784 and 1R01 HL31809, and by a grant from The Kroc Foundation.

### References

1. Payan, D. G., C. A. Hess, and E. J. Goetzl. 1984. Inhibition of somatostatin of the proliferation of T-lymphocytes and Molt-4 lymphoblasts. *Cell. Immunol.* 84:433-438.
2. Hinterberger, W., C. Cerny, H. Kinast, H. Pointer, and K. H. Tragl. 1977. Somatostatin reduces the release of colony-stimulating activity (CSA) from PHA-activated mouse spleen lymphocytes. *Experientia.* 34:860-862.
3. Payan, D. G., D. R. Brewster, and E. J. Goetzl. 1983. Specific stimulation of human T-lymphocytes by substance P. *J. Immunol.* 131:1613-1615.
4. Gilman, S. C., J. M. Schwartz, R. J. Milner, F. E. Bloom, and J. D. Feldman. 1982.  $\beta$ -Endorphin enhances lymphocyte proliferative responses. *Proc. Natl. Acad. Sci. USA.* 79:4226-4230.
5. Johnson, H. M., E. M. Smith, B. A. Torres, and J. E. Blalock. 1982. Regulation of the in vitro antibody response by neuroendocrine hormones. *Proc. Natl. Acad. Sci. USA.* 79:4171-4174.
6. Mathews, P. M., C. J. Froelich, W. L. Sibbitt, and A. D. Bankhurst. 1983. Enhancement of natural cytotoxicity by  $\beta$ -endorphin. *J. Immunol.* 130:1658-1662.
7. Wybran, J., T. Appelboom, J.-P. Famaey, and A. Govaerts. 1979. Suggestive evidence for receptors for morphine and methionine-enkephalin on normal human blood T-lymphocytes. *J. Immunol.* 123:1068-1070.
8. Beed, E. A., M. S. O'Dorisio, T. M. O'Dorisio, and T. S. Gaginella. 1983. Demonstration of a functional receptor for vasoactive intestinal peptide on Molt 4b T-lymphoblasts. *Regul. Pept.* 6:1-12.
9. Danek, A., M. S. O'Dorisio, T. M. O'Dorisio, and J. M. George. 1983. Specific binding sites for vasoactive intestinal poly-peptide on nonadherent peripheral blood lymphocytes. *J. Immunol.* 131:1173-1177.
10. Pernow, B. 1983. Substance P. *Pharmacol. Rev.* 35:85-141.
11. Hokfelt, T., J. O. Kellerth, G. Nilsson, and B. Pernow. 1975. Substance P: localization in the central nervous system and in some primary sensory neurons. *Science (Wash. DC).* 190:889-890.
12. Costa, M., J. B. Furness, R. Franco, I. Llewellyn-Smith, R. Murphy, and A. M. Beardsley. 1982. Substance P in nerve tissue in the gut. *Ciba Found. Symp.* 91:129-144.
13. Cuello, A. C., and I. Kanazawa. 1978. The distribution of substance P immunoreactive fibers in the rat central nervous system. *J. Comp. Neurol.* 178:129-156.
14. Johnson, A. R., and E. G. Erdos. 1973. Release of histamine from mast cells by vasoactive peptides. *Proc. Soc. Exp. Biol. Med.* 142:1252-1256.
15. Lembeck, F., J. Donnerer, and F. C. Colpaert. 1981. Increase of substance P in primary afferent nerves during chronic pain. *Neuropeptides.* 1:175-180.
16. Ljungdahl, A., T. Hokfelt, and G. Nilsson. 1978. Distribution of substance P-like immunoreactivity in the central nervous system of the rat. *Neuroscience.* 3:861-976.
17. Bar-Shavit, Z., R. Goldman, Y. Stabinsky, P. Gottlieb, M. Fridkin, V. I. Teichberg, and S. Blumberg. 1980. Enhancement of phagocytosis—a newly found activity of substance P residing in its N-terminal tetrapeptide sequence. *Biochem. Biophys. Res. Comm.* 94:1445-1451.
18. Lembeck, F., R. Gamse, and H. Juan. 1977. Substance P and sensory nerve endings. In Substance P. U. S. von Euler and B. Pernow, editors. Raven Press, New York. 169-181.
19. Hartung, H. P., and K. V. Toyka. 1983. Activation of macrophages by Substance P: induction of oxidative burst and thromboxane release. *Eur. J. Pharm.* 89:301-305.
20. Hanley, M. R. 1982. Substance P antagonists. *Trends Neurosci.* 5:138-139.
21. Payan, D. G., D. E. Trentham, and E. J. Goetzl. 1982. Modulation of human lymphocyte function by C3a and C3a(70-77). *J. Exp. Med.* 156:756-765.
22. Valone, F. H., E. Coles, V. R. Reinhold, and E. J. Goetzl. 1982. Specific binding of phospholipid platelet-activating factor by human platelets. *J. Immunol.* 129:1637-1641.
23. Payan, D. G., and E. J. Goetzl. 1983. Specific suppression of human T-lymphocyte function by leukotriene B<sub>4</sub>. *J. Immunol.* 131:551-553.
24. Bar-Shavit, Z., S. L. Teitelbaum, P. Reitsma, A. Hall, L. E. Pegg, J. Trial, and A. J. Kahn. 1983. Induction of monocytic differentiation and bone resorption by 1,25-dihydroxyvitamin D<sub>3</sub>. *Proc. Natl. Acad. Sci. USA.* 80:5907-5911.
25. Mangelsdorf, D. J., H. P. Koeffler, C. A. Donaldson, J. W. Pike, and M. R. Hanssler. 1984. 1,25-Dihydroxyvitamin D<sub>3</sub>-induced differentiation in human promyelocytic leukemia cell line (HL-60): receptor-mediated maturation to macrophage-like cells. *J. Cell Biol.* 98:391-398.
26. Watt, K. W. K., I. L. Brightman, and E. J. Goetzl. 1983. Isolation of two polypeptides comprising the neutrophil-immobilizing factor of human leukocytes. *Immunology.* 48:79-86.
27. Goldman, D. W., and E. J. Goetzl. 1982. Specific binding of leukotriene B<sub>4</sub> to receptors on human polymorphonuclear leukocytes. *J. Immunol.* 129:1600-1604.
28. Nawa, H., T. Hirose, H. Takashima, S. Inayama, and S. Nakanishi. 1983. Nucleotide sequences of cloned cDNAs for two types of bovine brain substance P precursor. *Nature (Lond.).* 306:32-36.
29. Murakoshi, T., M. Yanagisawa, C. Kitada, M. Fujino, and M. Otsuka. 1983. The role of the N-terminus in the active conformation of the substance P analogues. *Eur. J. Pharm.* 90:133-137.
30. Payan, D. G., A. Missirlian-Bastian, and E. J. Goetzl. 1984. Human T-lymphocyte subset specificity of the regulatory effects of leukotriene B<sub>4</sub>. *Proc. Natl. Acad. Sci. USA.* 81:3501-3505.

# Regulation of Calcium Homeostasis in Sensory Neurons by Bradykinin

Stanley A. Thayer, Teresa M. Perney, and Richard J. Miller

Department of Pharmacological and Physiological Sciences, University of Chicago, Chicago, Illinois 60637

The nonapeptide bradykinin (BK) activates sensory neurons and stimulates the transmission of nociceptive information into the CNS. We investigated the effect of this peptide on rat dorsal root ganglion neurons (DRG) grown *in vitro*. BK stimulated the synthesis of inositol trisphosphate ( $IP_3$ ) and the breakdown of phosphatidylinositol bisphosphate, the synthesis of diacylglycerol, and the release of arachidonic acid from DRG cells. The release of  $IP_3$  and arachidonic acid was not inhibited by pretreatment of the cells with pertussis toxin. BK also mobilized intracellular  $Ca^{2+}$  stores in DRG cells as assessed by fura-2-based microfluorimetry. Two types of  $Ca^{2+}$  stores appeared to exist in DRG neurons. One type could be mobilized by caffeine ( $10^{-2}$  M), and this effect could be blocked by ryanodine in a use-dependent manner. These stores occurred primarily in the cell soma and were virtually absent from cell processes. A second type of store could be mobilized by BK, presumably through the mediation of  $IP_3$ . These latter stores were distributed equally between the cell soma and processes. Experiments with combinations of caffeine and BK suggested that the stores mobilized by these 2 agents may be separate entities. Both the caffeine and BK sensitive  $Ca^{2+}$  storage sites appeared to participate in buffering a  $Ca^{2+}$  load induced in DRG neurons by cell depolarization. The relevance of these observations to the mechanism of action of BK on sensory neurons is discussed.

The nonapeptide bradykinin (BK) is released from its precursors the kininogens by the action of the enzyme kallikrein in response to trauma (Erdos, 1979). BK production is one of a number of responses to injury initiated by the activation of Hagemann factor—other examples being the blood clotting and complement fixation cascades. Receptors for BK exist in a number of tissues, including the nervous system (Manning and Snyder, 1983), gastrointestinal and vascular smooth muscle (Manning et al., 1986; Beny et al., 1987), the gastrointestinal mucosa (Manning et al., 1982) and many others. When such receptors are activated, a multitude of responses are initiated which help to orchestrate the body's defenses against injury.

BK powerfully activates sensory neurons (Higashi et al., 1982), leading to the increased release of neurotransmitters such as substance P (Yaksh and Hammond, 1982). These neurotrans-

mitters mediate the transmission of nociceptive information into the spinal cord. Although such a role for BK is well established, the molecular mechanisms underlying these actions are obscure. In non-neuronal tissue, BK receptor activation leads to the stimulation of both phospholipase C and phospholipase  $A_2$  and the subsequent release of a host of powerful lipid-derived second-messenger molecules (Hong and Deykin, 1982; Miller, 1987b). These include inositol trisphosphate ( $IP_3$ ), diacylglycerol (DAG), and arachidonic acid (Miller, 1987b). These substances can now initiate  $Ca^{2+}$  mobilization, activation of protein kinase C (PKC), and eicosanoid biosynthesis, which can, in turn, lead to the synthesis of further second messengers. Studies on neuronal clonal cell lines have indicated that BK can also stimulate the production of lipid-derived intermediates in neuronal tissues (Yano et al., 1984; Francel and Dawson, 1986; Francel et al., 1987; Jackson et al., 1987; Miller, 1987b; Van Calcar and Heumann, 1987). Furthermore, various electrophysiological effects of BK and of lipid-derived second messengers have been demonstrated both in neuronal cell lines and authentic sensory neurons (Baccaglini and Hogan, 1983; Fowler et al., 1985; Reiser and Hamprecht, 1985; Higashida and Brown, 1986a, b, 1987; Higashida et al., 1986; Osugi et al., 1986a, b; Rang and Ritchie, 1987; Weinreich, 1986).

The release of neurotransmitters from sensory neurons can be triggered by increases in  $[Ca^{2+}]_i$ . In dorsal root ganglion (DRG) and other nerve cells, the resting  $[Ca^{2+}]_i$  is normally very low ( $\approx 10^{-7}$  M) (Miller, 1987a).  $[Ca^{2+}]_i$  can potentially increase either through  $Ca^{2+}$  influx from the cell exterior via a variety of membrane channels or by mobilization from intracellular  $Ca^{2+}$  stores (Miller, 1987a). In the present series of studies, we have examined the effects of BK on lipid metabolism and authentic sensory neurons *in vitro*. We find that BK activates phospholipases C and  $A_2$  and also modulates intracellular  $Ca^{2+}$  stores. Furthermore, sensory neurons seem to possess a  $Ca^{2+}$  store that can be mobilized by methylxanthines such as caffeine. The properties of the BK and caffeine-sensitive stores are compared.

## Materials and Methods

**Cell culture.** Pure populations of DRG neurons were cultured as described by Perney et al. (1986). Briefly, DRG neurons were dissected from thoracic and lumbar segments of 1- to 3-d-old Sprague-Dawley rats, incubated for 15 min at 37°C in collagenase/dispase (0.8 and 6.4 units/ml), and then dissociated into single cells by trituration through a Pasteur pipette. The cells were then plated on laminin-fibronectin-coated coverglasses (no. 1, 25 mm diameter). Cells were fed every 2-3 d with Ham's nutrient mixture F-12 supplemented with 5% heat-inactivated rat serum, 4% 17-d embryonic rat extract, 50 ng/ml NGF, 44 mM glucose, 2 mM L-glutamine, 1% MEM 100 × vitamins (GIBCO), and penicillin/streptomycin (100 units/ml and 100 µg/ml, respectively; GIBCO). Cultures were maintained at 37°C in a water-saturated atmosphere with 5% carbon dioxide. Sympathetic neurons were cultured

Received Oct. 19, 1987; revised Jan. 28, 1988; accepted Mar. 22, 1988.

Supported by PHS Grants DA-02121, DA-02575, and MH-40165 and by grants from Miles Pharmaceuticals and Marion Labs. T.M.P. was supported by PHS Training Grant GM-07151. S.A.T. was supported by F32 NS-08009.

Correspondence should be addressed to Prof. Richard J. Miller, Ph.D., Department of Pharmacological and Physiological Sciences, University of Chicago, 947 E. 58th Street, Chicago, IL 60637.

Copyright © 1988 Society for Neuroscience 0270-6474/88/114089-09\$02.00/0



from rat superior cervical ganglia (Perney et al., 1986) and central neurons from rat striatum (Murphy et al., 1987). Cultures were treated with  $10^{-5}$  M cytosine arabinoside for 48 hr, 12 hr after plating to suppress the growth of non-neuronal cells.

**Measurement of  $[\text{Ca}^{2+}]_i$ .**  $[\text{Ca}^{2+}]_i$  was determined using a microfluorimeter to monitor  $\text{Ca}^{2+}$ -sensitive fluorescent chelator, fura-2 (Grynkiewicz et al., 1985). Neurons were loaded with the dye by incubation with 2  $\mu\text{M}$  fura-2 acetoxymethyl ester (Molecular Probes Inc., Eugene, OR), which is membrane permeant, for 1 hr at  $37^\circ\text{C}$  in HEPES-buffered Hank's balanced salt solution, pH 7.45, containing 0.5% BSA. The HEPES Hank's solution was composed of (in mM) the following: HEPES, 20; NaCl, 137;  $\text{CaCl}_2$ , 1.3;  $\text{MgSO}_4$ , 0.4;  $\text{MgCl}_2$ , 0.5; KCl, 5.0;  $\text{KH}_2\text{PO}_4$ , 0.4;  $\text{NaHPO}_4$ , 0.6;  $\text{NaHCO}_3$ , 3.0; and glucose, 5.6. Following the loading incubation, during which time the dye ester is hydrolyzed by cytosolic esterases to the membrane-impermeant polycarboxylate anion that is fura-2, the cells were washed twice in the HEPES-Hank's solution and incubated for 30 min. It is difficult to measure precisely the amount of dye loaded into a cell. However, in cells loaded with a 100  $\mu\text{M}$  fura-2 via internal dialysis with a patch-clamp pipette, the fluorescence signal was similar to the very brightest cells used in this study. Thus, 100  $\mu\text{M}$  would be an upper limit for the intracellular fura-2 concentration.

The coverglasses containing the loaded and washed cells were then mounted in a flow-through chamber for viewing. Briefly, the chamber consisted of a Plexiglas block machined to accommodate the coverslip as a bottom. Three reservoirs were cut into the block such that a thin sheet of buffer flowed out of the inlet reservoir, across the cells in the experimental chamber, and was drawn up across nylon mesh to the efflux reservoir for evacuation by suction. Solutions in the chamber could be completely exchanged within 15 sec, including the time taken to flow through the tubing delay between the large media reservoirs and the inlet to the chamber. Experiments were run at  $22^\circ\text{C}$ .

The perfusion chamber was mounted on an inverted microscope (Diaphot, Nikon, Garden City, NY), and cells and processes were localized by standard phase-contrast illumination. The perfusion chamber as well as the microfluorimeter are described in detail elsewhere (Thayer et al., 1987a, 1988). For excitation of the fura-2, the collimated light beam from a 200 W Hg arc lamp was passed through a dual-beam spectrophotometer (Phoenix Instruments), which alternated wavelengths from 340 to 380 nm by means of a wheel spinning at a frequency of 60 Hz. In place of the original sample chamber, a collimating beam probe was placed for focusing the light onto the end of a liquid guide (3 mm  $\times$  1 m, Oriol, Stratford, CT). On the other end of the liquid light guide, a similar probe was positioned for directing light through the epifluorescence illuminator of the microscope. The light was reflected off a dichroic mirror (Nikon, DM 400) and through a  $\times 70$  phase-contrast oil-immersion objective (E. Leitz Inc., Rockleigh, NJ, numerical aperture, 1.15). The emission fluorescence was selected for wavelength with a 480 nm barrier filter, and recordings were defined spatially with a rectangular diaphragm. The fluorescence emission was analyzed with a photomultiplier tube and discriminator (Thorn EMI Gencom Inc., Plainview, NY). The discriminator output was converted to pulses, which were then integrated by passing the signal through an 8-pole low-pass Bessel filter at 500 Hz. The photomultiplier signal was fed into one channel of an analog-to-digital converter computer system (C-lab, Indec Systems, Sunnyvale, CA). The signals from 2 photodiodes, each placed in a small portion of the light beam directed to the monochromators, were fed into 2 additional channels of the analog-to-digital converter.

The photomultiplier output was sorted into signal from 340 and 380 nm excitation by using the photodiode output as synchronizing signals. In the typical traces described here, 30 ratios were determined per second, the average ratio was displayed on-line, and the average intensity values for each wavelength were stored. After completion of a given experiment, the microscope stage was adjusted so that no cells or debris occupied the field of view defined by the diaphragm, and then background light levels were determined. Background light levels ranged from less than 5% for large cell bodies to close to 50% for very fine processes. Autofluorescence from cells that had not been loaded with fura-2 was not detectable. Records were later corrected for background and the ratios recalculated. Ratios were converted to free  $[\text{Ca}^{2+}]_i$  by using the equation  $[\text{Ca}^{2+}]_i = K(R - R_{\min})/(R_{\max} - R)$ , in which  $R$  is the 340/380 nm fluorescence ratio (Grynkiewicz et al., 1985). The maximum ratio ( $R_{\max}$ ), the minimum ratio ( $R_{\min}$ ), and the constant  $K$  ( $K$  is the product of the dissociation constant for fura-2 and the ratio of the free and bound forms of the dye at 380 nm) were determined from a standard curve to which the above equation was fit using a nonlinear least-squares

fit computer program. The system was recalibrated following any adjustment in the apparatus. Values for the constants  $R_{\min}$ ,  $R_{\max}$ , and  $K$  ranged from 0.121 to 0.334, 4.02 to 5.06, and 2034 to 2373, respectively. The standard curve was determined from the fura-2 pentapotassium salt in calibration buffer (which contains, in mM, HEPES, 20; KCl, 120; NaCl, 5; pH 7.1) containing 10 mM EGTA,  $K_d = 3.969 \times 10^6$  M (Fabiato and Fabiato, 1979), and varying amounts of added  $\text{Ca}^{2+}$ , which were calculated to give free  $\text{Ca}^{2+}$  concentrations ranging from 0 to 2000 nM. Records were digitally filtered with an algorithm that added  $\frac{1}{2}$  the value of each data point with  $\frac{1}{4}$  of the value of each of the 2 neighboring points. The data were cycled through this routine 5 times. Results are presented as means  $\pm$  SEM.

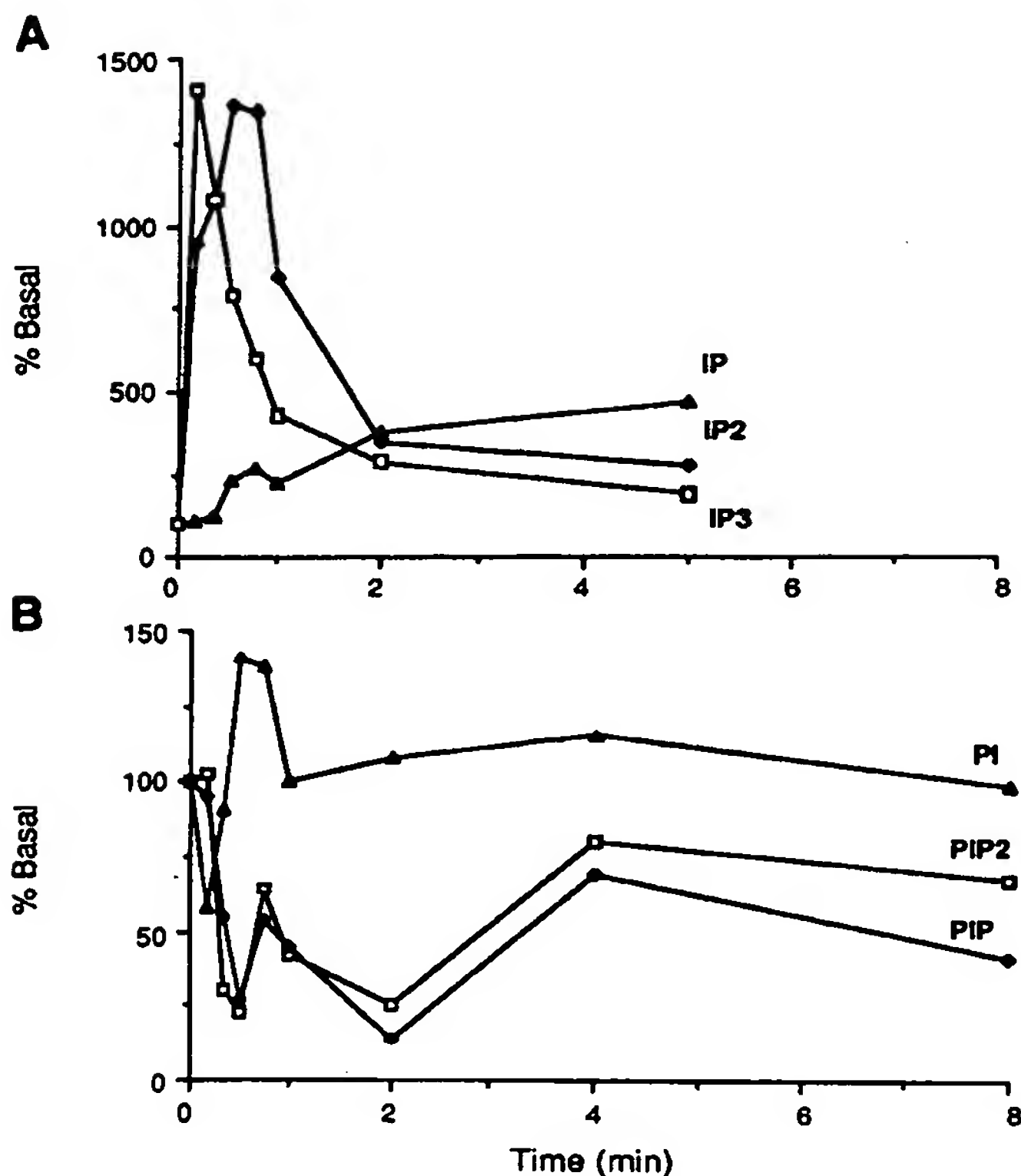
All experiments were performed on cells continuously perfused with HEPES Hank's solution. Depolarization-induced  $\text{Ca}^{2+}$  influx was produced by changing the perfusing solution from low  $\text{K}^+$  (5 mM) to high  $\text{K}^+$  (50 mM) with  $\text{K}^+$  exchanged for  $\text{Na}^+$  reciprocally.  $\text{Ca}^{2+}$ -free solutions were prepared by replacing  $\text{Ca}^{2+}$  with 20  $\mu\text{M}$  EGTA.

**Lipid metabolism.** DRG cultures were incubated with either  $^3\text{H}$ -myo-inositol (5  $\mu\text{Ci}/\text{ml}$ ) for 48 hr or  $^3\text{H}$ -arachidonic acid (1  $\mu\text{Ci}/\text{ml}$ ) for 10 hr in normal growth medium. In some experiments, cells were also incubated with 100–350 ng/ml pertussis toxin for 24 hr. Unincorporated isotope was removed by washing cultures 2 times with 2.4 ml of HEPES-buffered Eagle's MEM containing 0.5 mg/ml fatty acid-free BSA. Cells were then incubated for 30 min in 2.5 ml HEPES Eagle's MEM at  $37^\circ\text{C}$ . A final wash was performed, and the experiment was initiated after 5 min by the addition of BK (100 nM except as indicated). The reaction was terminated after various intervals by rapid aspiration of the buffer and addition of either 2 ml of 10% trichloroacetic acid (TCA) or 2 ml of methanol for  $^3\text{H}$ -myo-inositol- or  $^3\text{H}$ -arachidonic acid-labeled cells, respectively. In experiments in which arachidonic acid release was studied, the aspirated medium was saved. Cells were then scraped off the plate and added to 15 ml glass centrifuge tubes. The TCA extracts were spun at 1000 rpm for 5 min. The supernatant was saved for inositol phosphate analysis and the pellet for phospholipid analysis. Chloroform, 2 ml, and HCl, 100  $\mu\text{l}$ , were added to the methanol extracts and phase-separated by the addition of 1.2 ml of 1 N HCl/5 mM EGTA. The aqueous phase was removed, and the organic phase evaporated under  $\text{N}_2$  and saved for DAG analysis.

**Analysis of inositol sugars.**  $^3\text{H}$ -labeled inositol sugars were assayed by the method of Berridge et al. (1982). Briefly, TCA extracts containing the inositol sugars were extracted 5 times in an equal volume of ether to remove the acid and then neutralized with 50 mM tetraborate. The water-soluble inositol phosphates were then separated by ion-exchange chromatography on Dowex-1 X 8 (formate form). The samples were applied to 1 ml columns of Dowex, and free inositol was washed through the columns with ten 2 ml rinses of glass-distilled water. Subsequent elution of glycerol phosphoinositol, inositol monophosphate (IP), inositol bisphosphate (IP<sub>2</sub>), and IP<sub>3</sub> was accomplished by five 2 ml rinses with the following buffers: 60 mM sodium formate/5 mM sodium tetraborate; 0.1 M formic acid/0.2 M ammonium formate; 0.1 M formic acid/0.5 M ammonium formate; and 0.1 M formic acid/1.0 M ammonium formate.

**Phospholipid analysis.** TCA pellets were resuspended in 1.5 ml of chloroform/methanol/HCl (100:100:1, vol/vol), sonicated, and spun at 1000 rpm for 5 min. The supernatant was saved and the pellet extracted twice more in a similar manner first with 1.5 ml of chloroform/methanol/HCl (100:100:1) and then with 1 ml of chloroform/methanol/HCl (200:100:1). The supernatants were pooled and phase-separated by the addition of 1.5 ml chloroform and 1.5 ml 0.1 M HCl. The aqueous phase was removed and the organic phase washed twice with the theoretical upper phase [chloroform/methanol/water (3:48:47)] before drying under  $\text{N}_2$ . After evaporation the residues were redissolved in 50 ml of chloroform and applied to TLC plates (silica G). The phospholipids were separated on oxalate-impregnated silica plates and developed in chloroform, methanol, and 4 N  $\text{NH}_4\text{OH}$  (45:35:9.5) according to the method described by Billah and Lapetina (1982). Individual lipids were visualized by exposure to  $\text{I}_2$  vapors. Zones corresponding to lipid standards were scraped and assayed for radioactivity by liquid scintillation methods.

**DAG analysis.** For  $^3\text{H}$ -arachidonic acid-labeled DAG determination, the dried-down lipids were redissolved in 1 ml chloroform and applied to a 2 ml silicic acid column. Neutral lipids were eluted from the column by a 10 ml rinse with chloroform. After evaporation of the column eluate under  $\text{N}_2$ , the residue was redissolved in 50  $\mu\text{l}$  of chloroform and applied to TLC plates. Separation of DAG was achieved by developing



**Figure 1.** Time course of BK-stimulated phospholipid breakdown and inositol phosphate production. Cultured DRG neurons were prelabeled with  $^3\text{H}$ -myoinositol ( $5 \mu\text{Ci}/\text{ml}$ ) for 48 h and then exposed to BK ( $100 \text{ nM}$ ) for various intervals (see Materials and Methods). *A*, Inositol phosphate production. The data are expressed as percentage of the basal value. Each point represents the mean of duplicate determinations. This time course is representative of 3 experiments. The basal levels were  $^3\text{H}$ -IP<sub>3</sub>, 488 cpm/plate;  $^3\text{H}$ -IP<sub>2</sub>, 532 cpm/plate; and IP, 2448 cpm/plate. *B*, Phospholipid breakdown. The data are the mean of duplicate determinations and are expressed as percentage of the basal levels. Basal levels were PIP<sub>2</sub>,  $8.83 \times 10^3$  cpm/plate; PIP,  $8.75 \times 10^3$  cpm/plate; and PI,  $1.6 \times 10^5$  cpm/plate.

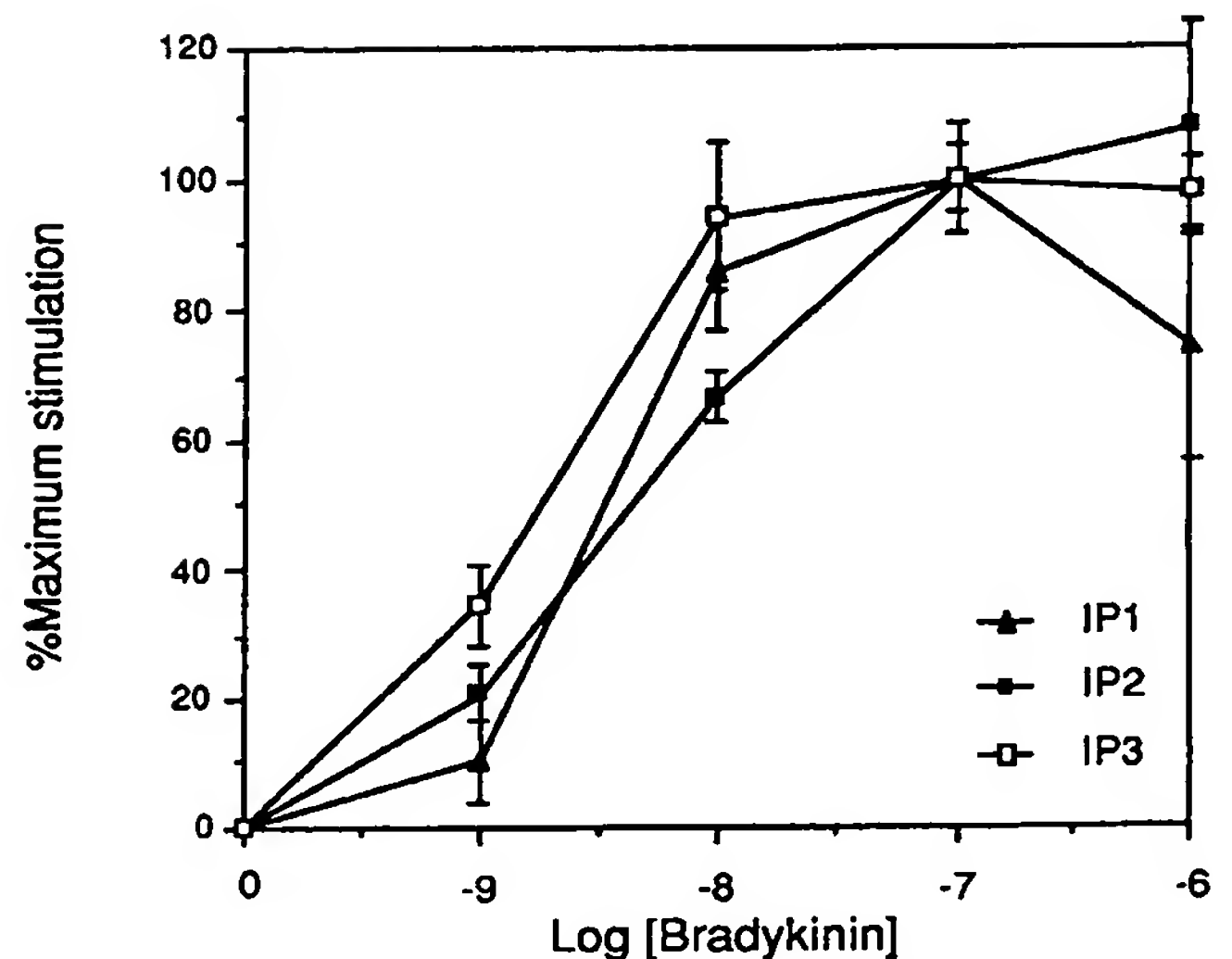
the plates in ether/hexane/acetic acid (70:30:1) as described by Griendling et al. (1986).

**Arachidonic acid release.** The release of  $^3\text{H}$ -arachidonic acid and its metabolites into the medium was determined by assaying 200  $\mu\text{l}$  aliquots of the medium for radioactivity by liquid scintillation spectrophotometry. In some experiments,  $^3\text{H}$ -arachidonic acid in the organic phase was separated from its metabolites by TLC [developed in chloroform/isopropanol/ethanol/formic acid (45:5:0.5:0.3)] after extraction of the medium in 6 ml chloroform/methanol (2:1). The 2 methods yielded similar results.

## Results

### Phospholipid metabolism

In many cell types, BK has been shown to activate both phospholipase C and phospholipase A<sub>2</sub> (Miller, 1987b). For example, this has been demonstrated in some neuronal clonal cell lines, although not in authentic DRG cells (Yano et al., 1984; Francel and Dawson, 1986; Francel et al., 1987; Jackson et al., 1987; Van Calcar and Heumann, 1987). We therefore began by examining the effects of BK on phospholipid metabolism in cultures of rat DRG neurons *in vitro*. BK, in the range  $10^{-9}$ – $10^{-6}$  M, rapidly stimulated the production of IP, IP<sub>2</sub>, and IP<sub>3</sub> (Figs.



**Figure 2.** Dose dependence of BK-stimulated inositol phosphate production. Cultured DRG neurons were prelabeled with  $^3\text{H}$ -myoinositol ( $5 \mu\text{Ci}/\text{ml}$ ) for 48 hr and then exposed to BK at various concentrations for 30 sec (see Materials and Methods). Data are means  $\pm$  SE of triplicate determinations and are expressed as percentage of the control value.

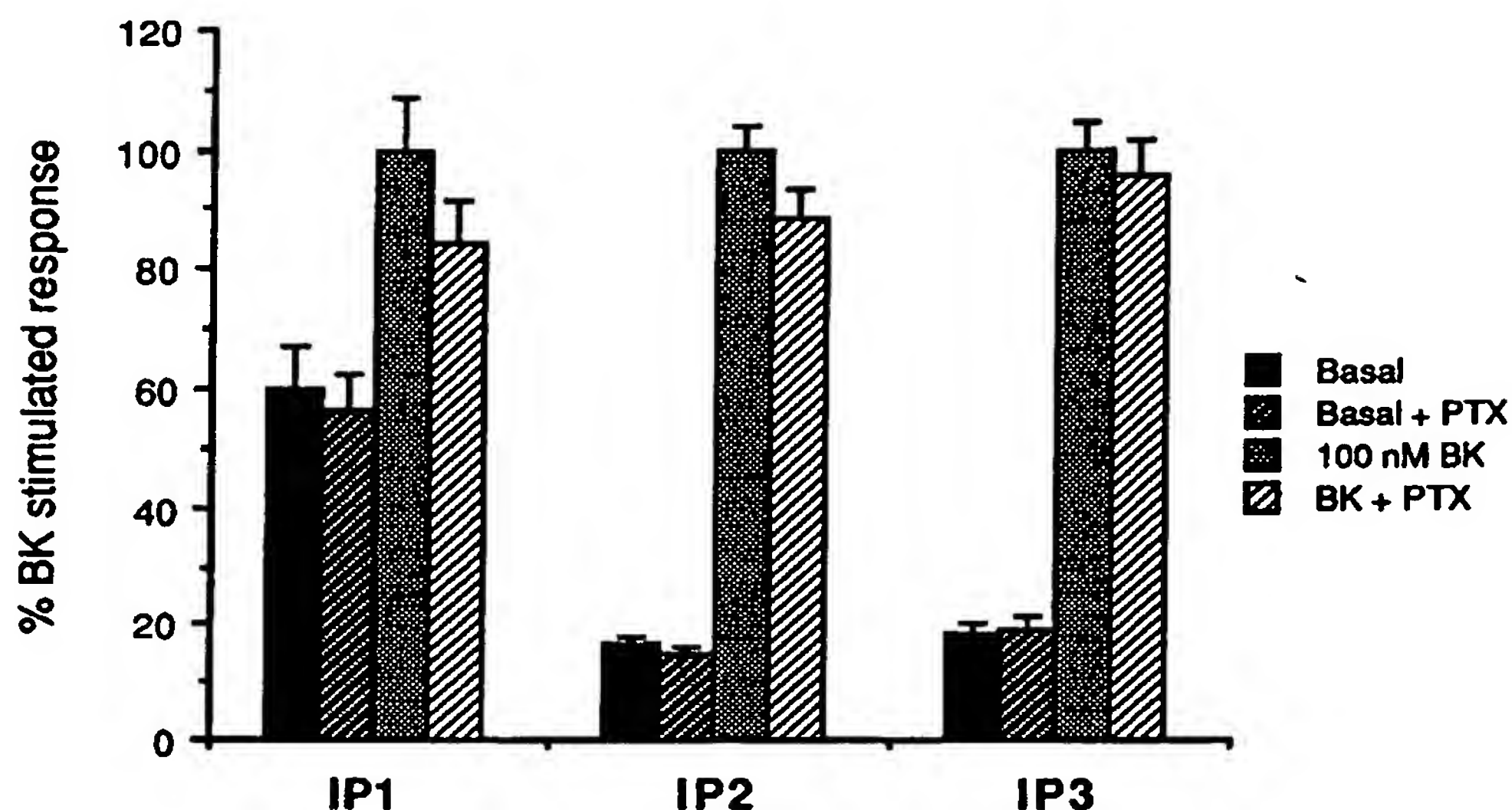
1 and 2). The synthesis of IP<sub>3</sub> was seen most rapidly followed by that of IP<sub>2</sub> and then IP. Large increases in the synthesis of IP<sub>3</sub> were observed (up to 20-fold). As reported for a variety of other systems, the agonist stimulated synthesis of IP<sub>3</sub> and IP<sub>2</sub> was relatively transient, declining to basal levels over a few minutes. In contrast, the slower and less profound synthesis of IP was more sustained and was still apparent at the termination of experiments (5 min). As would be expected during the period over which the synthesis of IP<sub>3</sub> and IP<sub>2</sub> occurred, the concentrations of PIP<sub>2</sub> and PIP fell (Fig. 1). BK-stimulated IP<sub>3</sub> production was not blocked after pretreatment of cells with pertussis toxin (PTX) (Fig. 3). BK also increased the production of DAG by DRG cells (Fig. 4). Interestingly, the time course of DAG production was considerably longer than that observed for IP<sub>3</sub>. An initial peak of DAG production occurred that correlated in time with the production of IP<sub>3</sub>. However, this was followed by a second extended peak of DAG production that lasted for many minutes. Such biphasic agonist-induced DAG production has now been noted in many cases and may reflect multiple sources of DAG (Miller, 1988). These observations make it clear that BK does activate phospholipase C in DRG neurons. We also found that BK activated phospholipase A<sub>2</sub>. BK stimulated the release of  $^3\text{H}$ -arachidonic acid and its metabolites from labeled DRG cells. In the presence of BK ( $10^{-7}$  M), the levels of radioactivity released into the culture medium rose to  $115.8 \pm 7.9\%$  of control after 30 sec and  $127.4 \pm 8.5\%$  of control after 5 min of incubation ( $n = 18$ ). This release was not significantly reduced by pertussis toxin treatment ( $n = 10$ ). In one experiment, we also observed that BK ( $10^{-7}$  M) stimulated the production of lyso-PI by the cells.

### Intracellular $\text{Ca}^{2+}$ stores

IP<sub>3</sub> has been shown to mobilize  $\text{Ca}^{2+}$  from intracellular stores in a number of cell types (Berridge, 1987). We therefore examined the effects of BK on  $[\text{Ca}^{2+}]$ , in different portions of rat DRG neurons and compared it with the effects of other stimuli

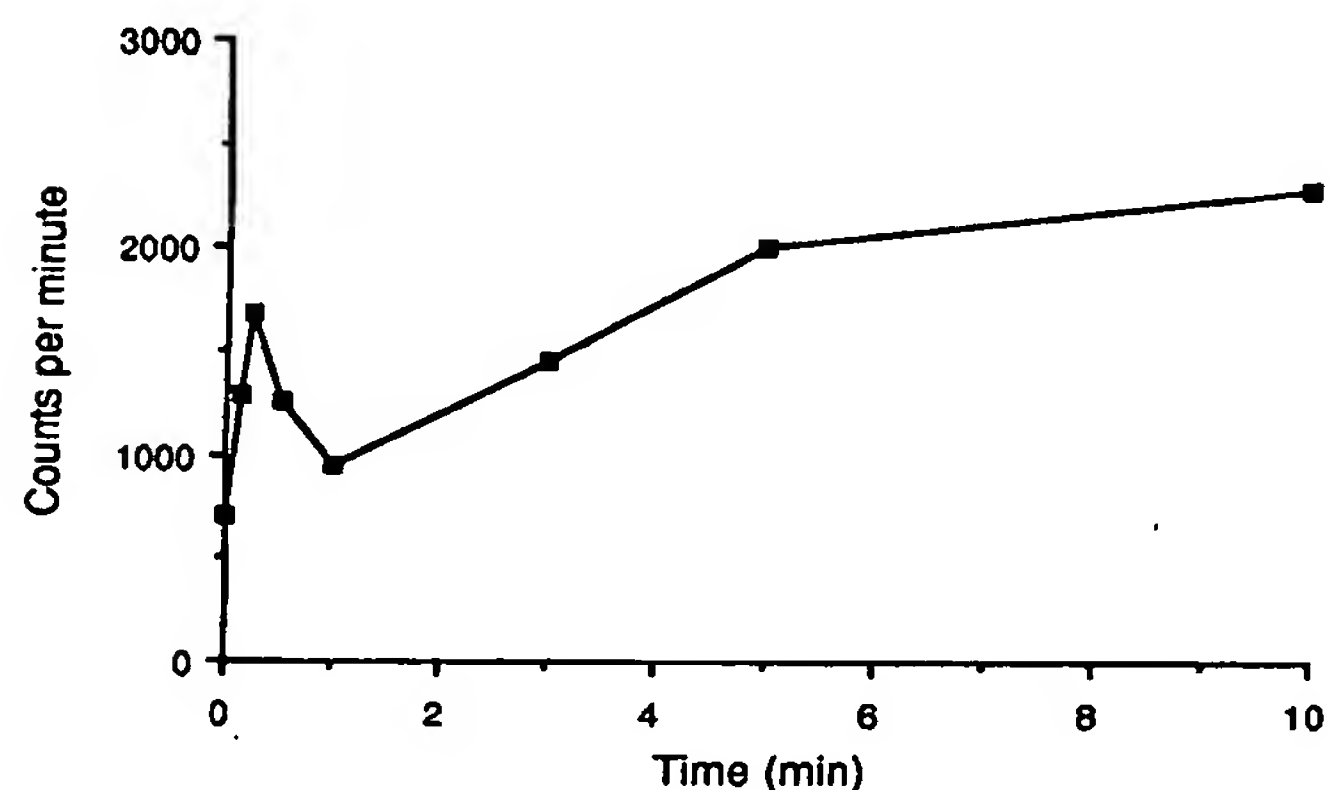


**Figure 3.** Effects of pertussis toxin on BK-stimulated inositol phosphate production. Cultured DRG neurons were prelabeled with  $^3\text{H}$ -myoinositol ( $5 \mu\text{Ci}/\text{ml}$ ) and then exposed to BK ( $100 \text{ nM}$ ) for 30 sec. For some of the cultures, 350 ng/ml pertussis toxin was added 24 hr after labeling had begun. Data are means  $\pm$  SE of triplicate determinations in 4 experiments and are expressed as percentage of the control value (see Materials and Methods). A similar experiment in which the time of exposure to BK was varied also showed no pertussis toxin sensitivity to the response.

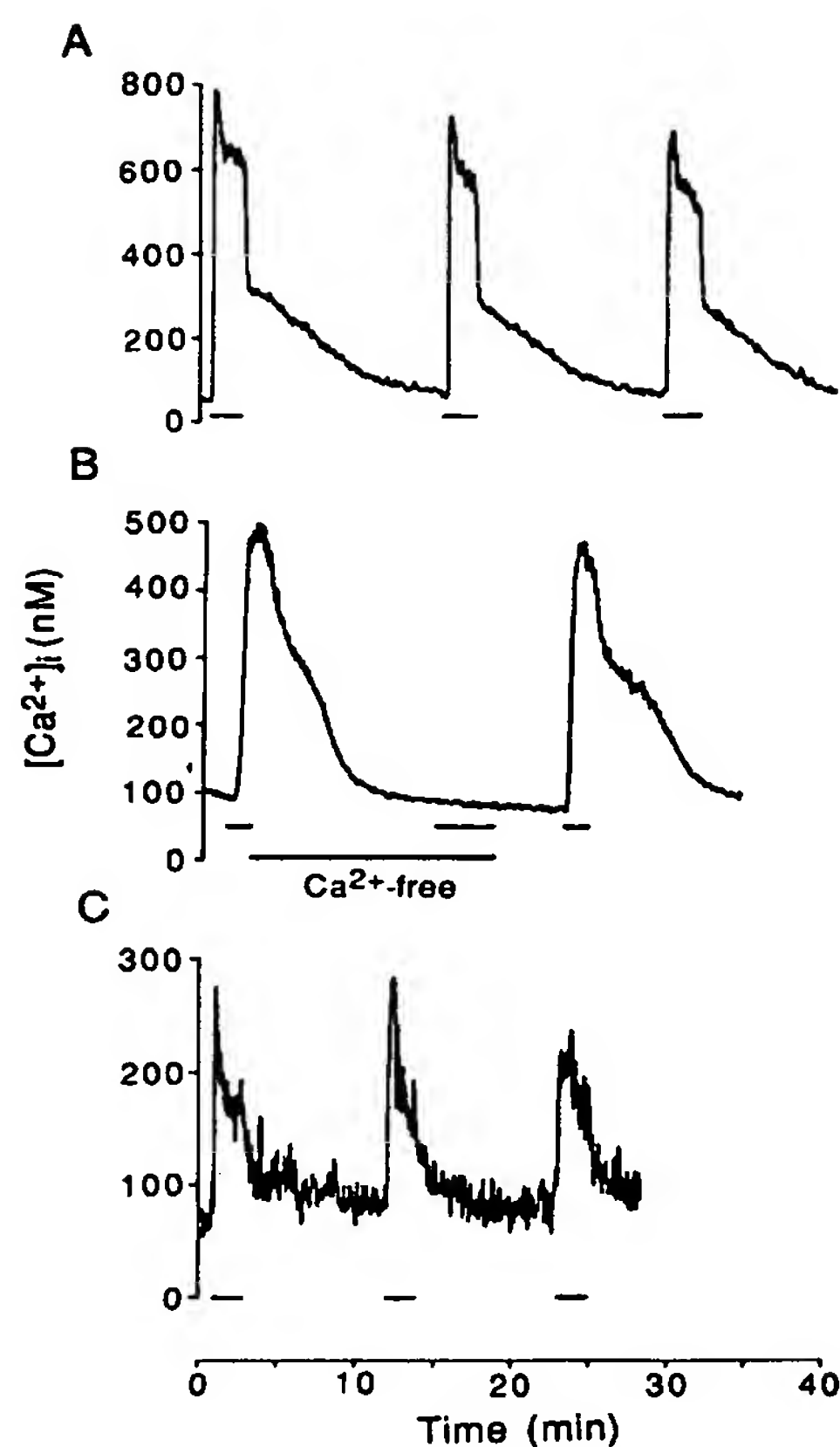


that affect  $[\text{Ca}^{2+}]_i$ . In the cell bodies of DRG cells, depolarization induced by raising  $[\text{K}^+]_o$  induced a rapid rise in  $[\text{Ca}^{2+}]_i$ . This occurred in 100% of the cells examined (Figs. 5, 9).  $[\text{Ca}^{2+}]_i$  rose from a resting concentration of  $77 \pm 7 \text{ nM}$  ( $n = 30$ ) to  $513 \pm 6 \text{ nM}$  ( $n = 27$ ) (Figs. 5, 9). In  $\text{Ca}^{2+}$ -free medium, the response to depolarization was completely blocked (Fig. 5B). A striking feature of the depolarization-induced  $\text{Ca}^{2+}$  transient in DRG cells was that it was buffered very slowly. Following washout of the depolarizing stimulus,  $[\text{Ca}^{2+}]_i$  initially fell rapidly but then stabilized and only declined very slowly even in  $\text{Ca}^{2+}$ -free medium (Figs. 5, 6). This slow buffering of  $\text{Ca}^{2+}$  stands in contrast to other types of peripheral and central neurons we have examined, which buffer similar  $\text{Ca}^{2+}$  loads much more quickly (Fig. 6) (Murphy et al., 1987; Thayer et al. 1987a).

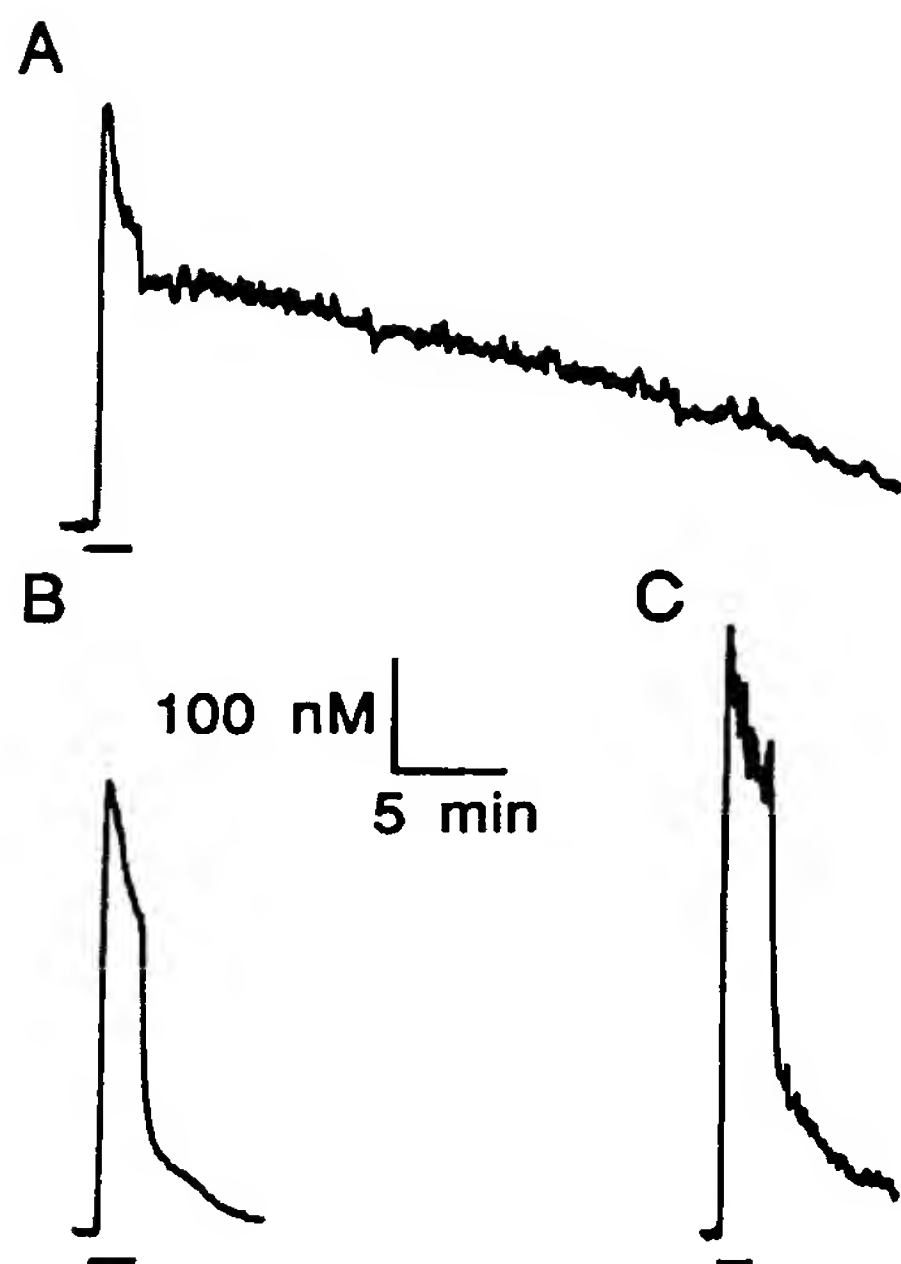
Methylxanthines such as caffeine have been shown to mobilize  $\text{Ca}^{2+}$  from intracellular stores in several types of cells, including peripheral neurons (Neering and McBurney, 1984; Lipscombe et al., 1987; Thayer et al., 1987b). Caffeine was also very effective in DRG cell bodies. In  $\text{Ca}^{2+}$ -free medium,  $10^{-2} \text{ M}$



**Figure 4.** Time course of bradykinin stimulated diacylglycerol formation. Cultured DRG neurons were prelabeled with  $^3\text{H}$ arachidonic acid ( $1 \mu\text{Ci}/\text{ml}$ ) for 10 hr and then exposed to bradykinin ( $100 \text{ nM}$ ) for various intervals. This time course is representative of three similar experiments. (See Materials and Methods).



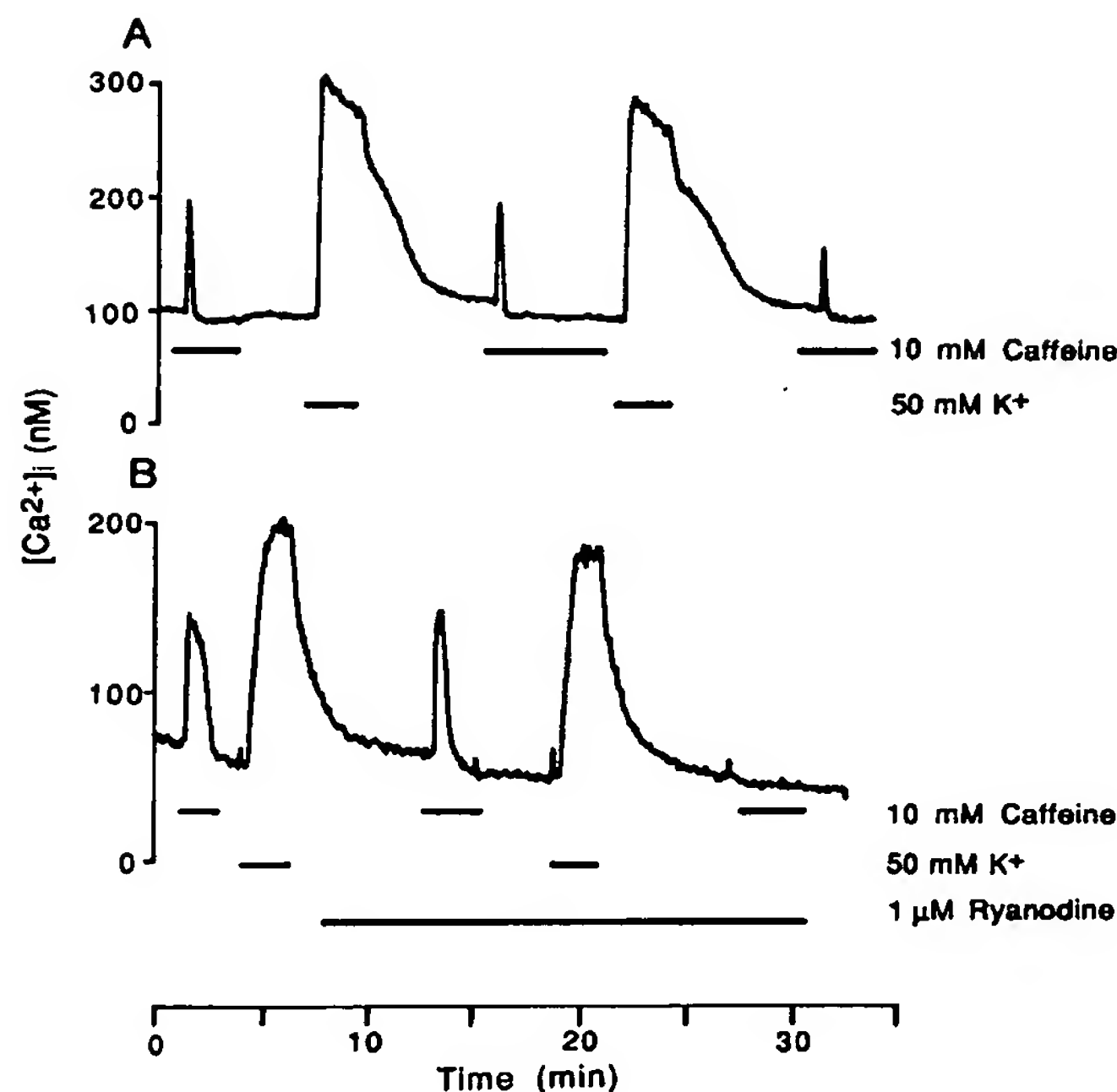
**Figure 5.** Depolarization-induced increases in  $[\text{Ca}^{2+}]_i$  in sensory neurons.  $[\text{Ca}^{2+}]_i$  was measured in single DRG cell bodies (A and B) and processes (C) as described in Materials and Methods. Cells were depolarized during the time indicated by the horizontal bars by changing the perfusing solution from low (5 mM) to high (50 mM)  $\text{K}^+$  media. In contrast to the 3 control responses generated in A, depolarization in  $\text{Ca}^{2+}$ -free ( $20 \mu\text{M}$  EGTA) media (horizontal bar) failed to elicit a response (B). Depolarization of a neuronal process (C) produced  $[\text{Ca}^{2+}]_i$  transients that were smaller and more rapidly buffered than those elicited in cell bodies.



**Figure 6.** Comparison of  $\text{Ca}^{2+}$  buffering in sensory, sympathetic, and central neurons. Single somata from the DRG (A), the SCG (B), and the striatum (C) were depolarized during the time indicated by the horizontal bars by changing the perfusing solution from low (5 mM) to high (50 mM)  $\text{K}^{+}$  media.  $[\text{Ca}^{2+}]_i$  was measured as described in Materials and Methods.

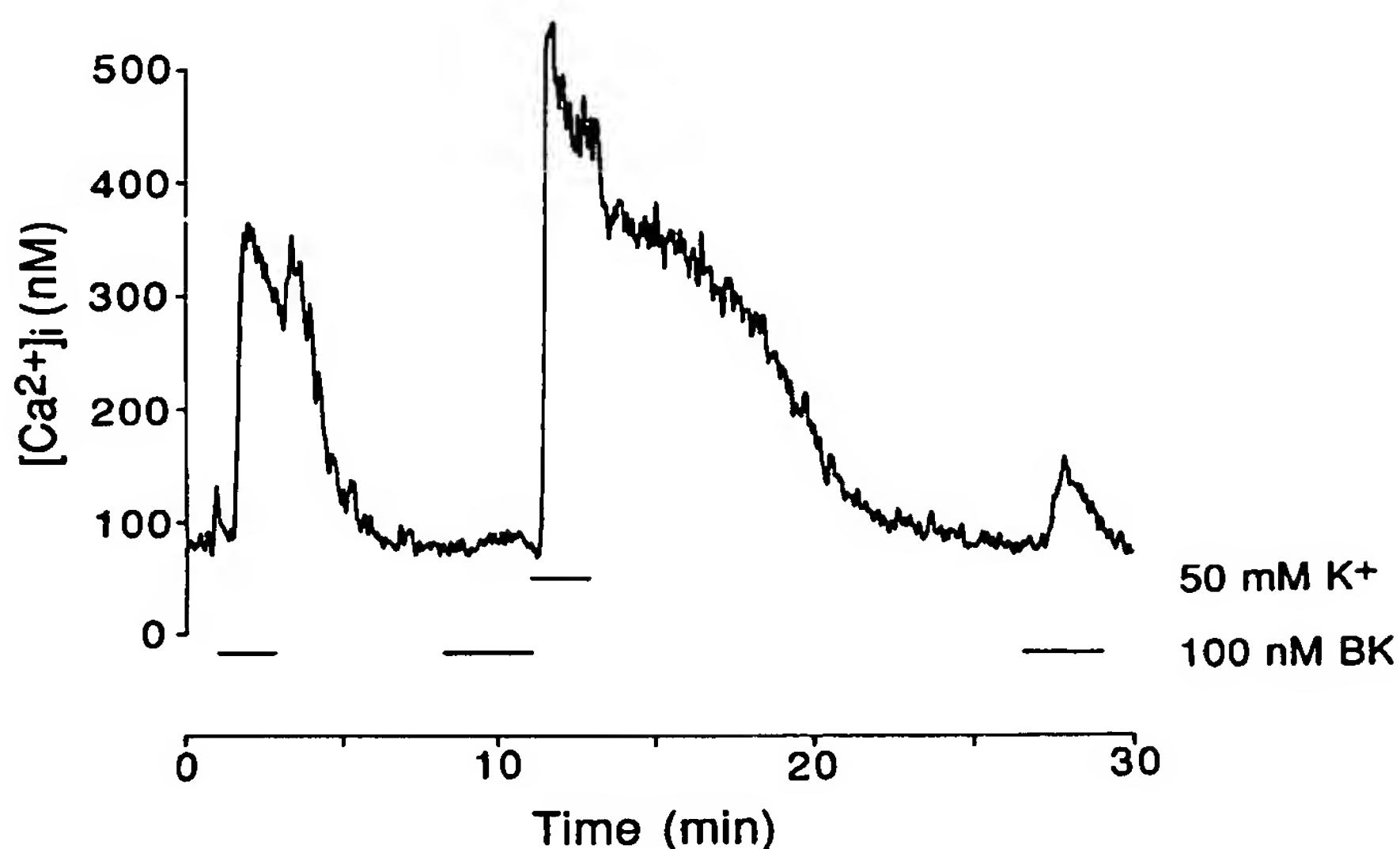
caffeine produced  $\text{Ca}^{2+}$  transients in the vast majority of cells examined (83%, Figs. 7, 9). In order to elicit multiple responses to caffeine, the stores required "refilling" after each exposure. This could be achieved by depolarizing the cell briefly in  $\text{Ca}^{2+}$ -containing medium and allowing  $\text{Ca}^{2+}$  influx to elevate  $[\text{Ca}^{2+}]_i$  (Fig. 7). The caffeine-sensitive stores in DRG neurons seem to be similar to those seen in muscle cells, as the caffeine-induced increases in  $[\text{Ca}^{2+}]_i$  could be completely blocked by ryanodine (Fig. 7). The block produced by ryanodine was "use dependent." Thus, after addition of ryanodine, a caffeine response could be obtained initially; however, after this first response, all subsequent attempts to elicit a caffeine response, even after normal "refilling," were completely blocked. We have observed that the ryanodine block of the caffeine response had precisely the same characteristics in rat sympathetic neurons (Thayer et al., 1987b). BK also produced  $\text{Ca}^{2+}$  transients in DRG cell bodies. These were still observed in  $\text{Ca}^{2+}$ -free medium. The magnitude of the  $\text{Ca}^{2+}$  transients produced by BK were similar to those produced by caffeine, although they were observed in considerably fewer cells (33%; Figs. 8, 9). The effects produced by BK desensitized. Following the production of a BK response, subsequent applications of the peptide produced progressively smaller responses even if attempts were made to refill stores as with caffeine (Fig. 8).

The effects of these various stimuli on  $[\text{Ca}^{2+}]_i$  in individual DRG cell processes were also examined. Several interesting differences were observed in comparison with results obtained in cell bodies. First, we found that elevating  $[\text{K}^{+}]_o$  was less effective in raising  $[\text{Ca}^{2+}]_i$  in cell processes (Figs. 5, 9). Although this depolarizing stimulus was still effective in every process examined, the net rise in  $[\text{Ca}^{2+}]_i$  to  $219 \pm 37 \text{ nM}$  ( $n = 23$ ) was



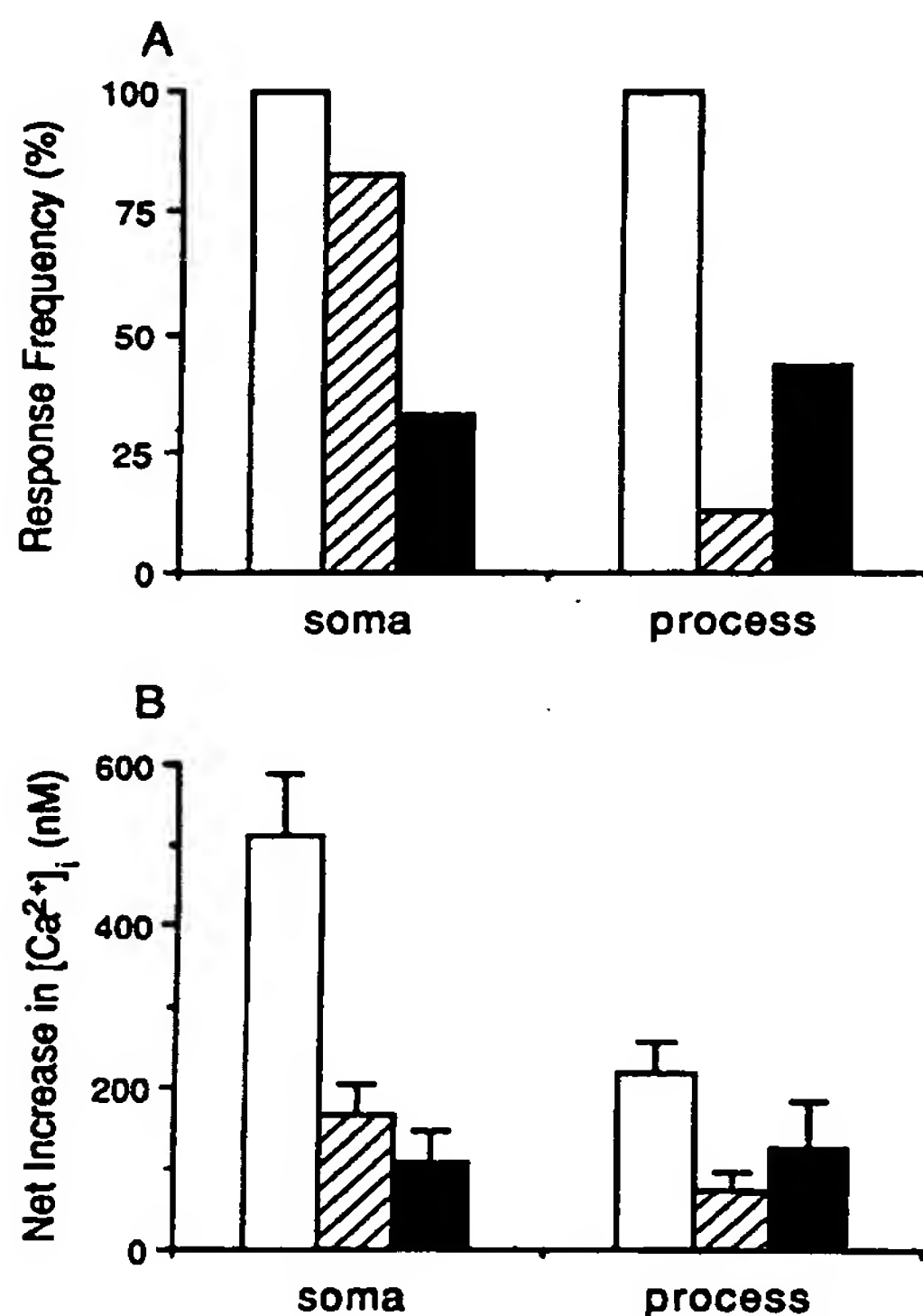
**Figure 7.** Caffeine-sensitive  $\text{Ca}^{2+}$  stores in sensory neurons. In DRG neurons perfused with  $\text{Ca}^{2+}$ -free (20  $\mu\text{M}$  EGTA) medium, adding 10 mM caffeine to the solution at the times indicated, produced a rapid and transient increase in  $[\text{Ca}^{2+}]_i$ . Multiple responses to caffeine could be elicited only if the intracellular stores were allowed to refill with  $\text{Ca}^{2+}$ , accomplished here by perfusing with depolarizing (50 mM  $\text{K}^{+}$ ) media containing normal  $\text{Ca}^{2+}$  levels. When 1  $\mu\text{M}$  ryanodine was added to the perfusing medium prior to and during the second application of caffeine (B), the response to the third application of caffeine was always blocked.  $[\text{Ca}^{2+}]_i$  was measured as described in Materials and Methods.

considerably smaller than that observed in the soma (Figs. 5, 9). This is in contrast to sympathetic neurons, for example, in which we observed that raising  $[\text{K}^{+}]_o$  increased  $[\text{Ca}^{2+}]_i$  in cell bodies and processes to a similar extent (Thayer et al., 1987a). The observations on DRG cells suggest some variability in the density of  $\text{Ca}^{2+}$  channels in different parts of the cell. In addition, increases in  $[\text{Ca}^{2+}]_i$  produced by depolarization in cell processes were much more transient than the long-lasting increases observed in cell bodies (Fig. 5). Striking differences were also observed with respect to the effects of caffeine in cell processes. Whereas caffeine was very effective in virtually all cell bodies tested, it rarely produced any effects in cell processes. We did observe occasional small responses to caffeine in processes, but they were very infrequent (Fig. 9). These observations are similar to others made in rat and bullfrog sympathetic neurons, in which the effects of caffeine on  $[\text{Ca}^{2+}]_i$  are also restricted to the cell soma (Lipscombe, 1987; Thayer et al., 1987a) (Figs. 9, 10). In contrast to caffeine, BK was at least as effective in cell processes as in the cell soma. The average increases in  $[\text{Ca}^{2+}]_i$  produced by BK in cell processes ( $126 \pm 57 \text{ nM}$ ,  $n = 10$ ) were similar to those observed in cell bodies ( $108 \pm 37 \text{ nM}$ ,  $n = 10$ ), and the response was seen more frequently (43 versus 33%; Fig. 9). These differential effects could indicate that the stores activated by caffeine and BK are different. That this may be so is further suggested by results such as those shown in Figure 10. There, we illustrate a cell body that failed to respond to BK but showed a large caffeine response. In contrast, a DRG process in



**Figure 8.** BK-induced increases in  $[\text{Ca}^{2+}]_i$ . The cell body of a sensory neuron was perfused with either 100 nM BK or depolarizing (50 mM  $\text{K}^+$ ) medium at the times indicated by the horizontal bars.  $[\text{Ca}^{2+}]_i$  was measured as described in Materials and Methods.

which caffeine failed to elicit a response showed a large BK effect. Such differential sensitivity was frequently observed. Although lack of a response to BK may indicate the absence of BK receptors in some instances, taken together these results strongly suggest that caffeine and BK ( $\text{IP}_3$ )-sensitive stores are separate entities.



**Figure 9.** Comparison of the increases in  $[\text{Ca}^{2+}]_i$  produced by depolarization, caffeine, and BK in somata and processes of sensory neurons. The frequency (A) and magnitude (B) of the responses in either DRG somata or processes to perfusion with media containing 50 mM  $\text{K}^+$  (open bars), 10 mM caffeine (hatched bars), or 100 nM BK (solid bars) are shown. Basal  $[\text{Ca}^{2+}]_i$  levels were subtracted from the peak responses and displayed as means  $\pm$  SE.

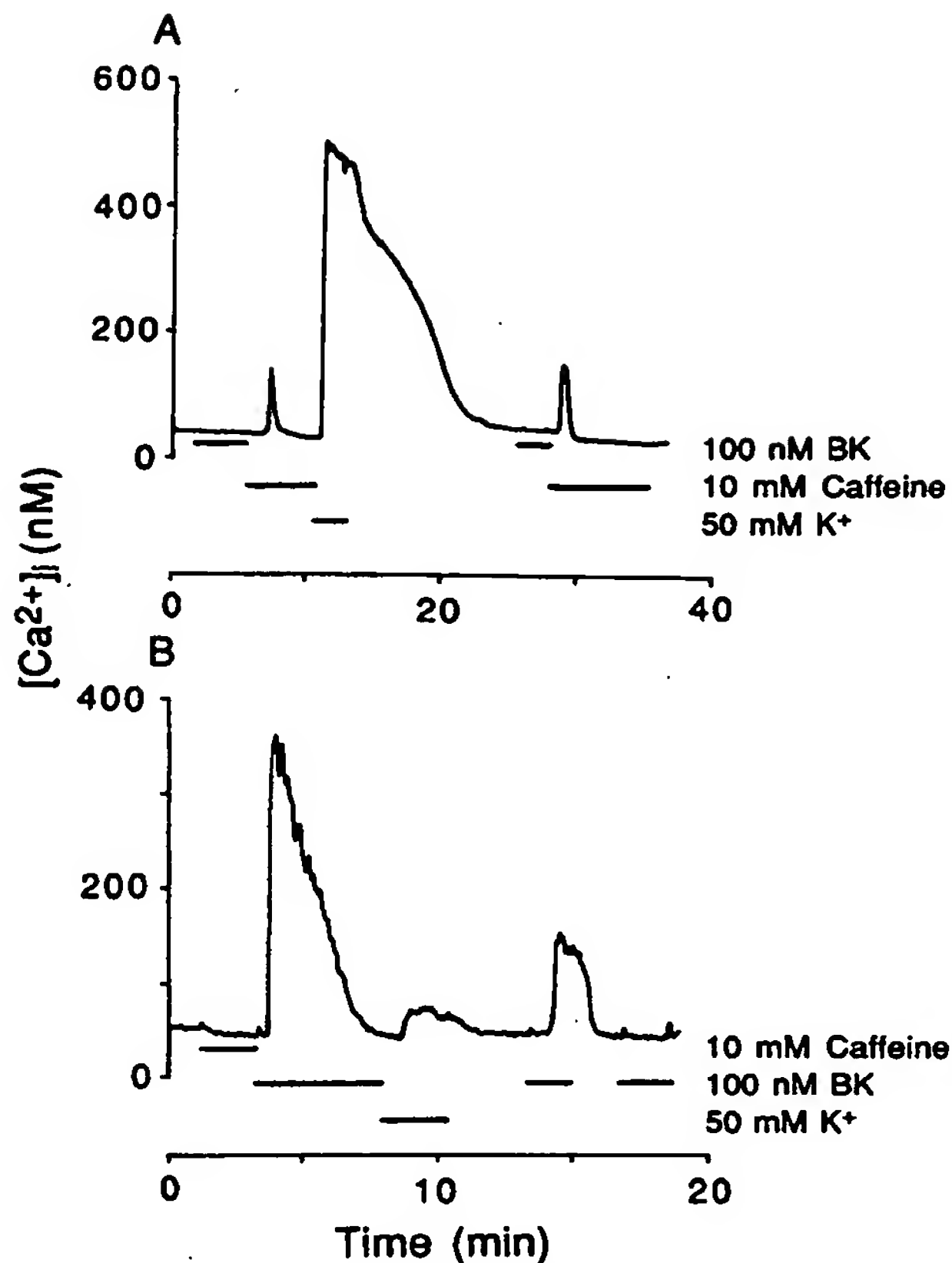
Another observation we have made concerns the potential role of intracellular  $\text{Ca}^{2+}$  storage sites in  $\text{Ca}^{2+}$  buffering. As discussed above, the shape of the  $\text{Ca}^{2+}$  transient evoked by depolarization in the cell bodies of DRG neurons is very characteristic. After a rapid initial decline, the  $\text{Ca}^{2+}$  signal decays only very slowly. We found, however, that the shape of this transient could be radically altered if either caffeine- or BK-sensitive stores were first discharged. After such treatment, the long-lasting tail of the  $\text{Ca}^{2+}$  transient was abolished or at least very greatly reduced (Fig. 11). Following washout of the caffeine or BK stimulus, the depolarization-evoked  $\text{Ca}^{2+}$  transients returned to their original form. Indeed, following discharge of intracellular stores, the  $\text{Ca}^{2+}$  transients obtained by depolarization of DRG cells resembled those obtained in other peripheral or central neurons (see above). We also observed that acute treatment of cells with phorbol esters or down-regulation of protein kinase C following chronic phorbol ester treatment (Matthies et al., 1987; Ewald et al., 1988) did not alter the buffering of  $\text{Ca}^{2+}$  by the cells. This indicates that the effects of BK on  $\text{Ca}^{2+}$  buffering do not involve protein kinase C but are presumably related to the discharge of  $\text{IP}_3$ -sensitive stores.

We have noted above that the activation of  $\text{IP}_3$  synthesis produced by BK in DRG neurons was not blocked by pertussis toxin. We therefore examined the effect of the toxin on BK-induced  $\text{Ca}^{2+}$  transients in cell processes. Quite surprisingly, we observed that, although BK was still effective in some instances, its ability to increase  $[\text{Ca}^{2+}]_i$  in pertussis toxin-treated cells seemed to be substantially decreased (13% of the processes responded after treatment,  $n = 19$  with a mean response of  $60 \pm 24$  nM;  $n = 3$ ). Caffeine, as noted, produced very little effect in cell processes; if anything, however, it was even less effective following the toxin (11% responded,  $n = 19$ ; mean 37 nM;  $n = 2$ ). In contrast, pertussis toxin had no effect upon the size of the  $\text{Ca}^{2+}$  transients produced in cell processes by depolarization (net increase  $261 \pm 47$  nM,  $n = 19$ ).

## Discussion

The action of BK on virtually all its target cells leads to the activation of both phospholipase C and phospholipase  $\text{A}_2$ . This, in turn, leads to the production of a variety of lipid-derived

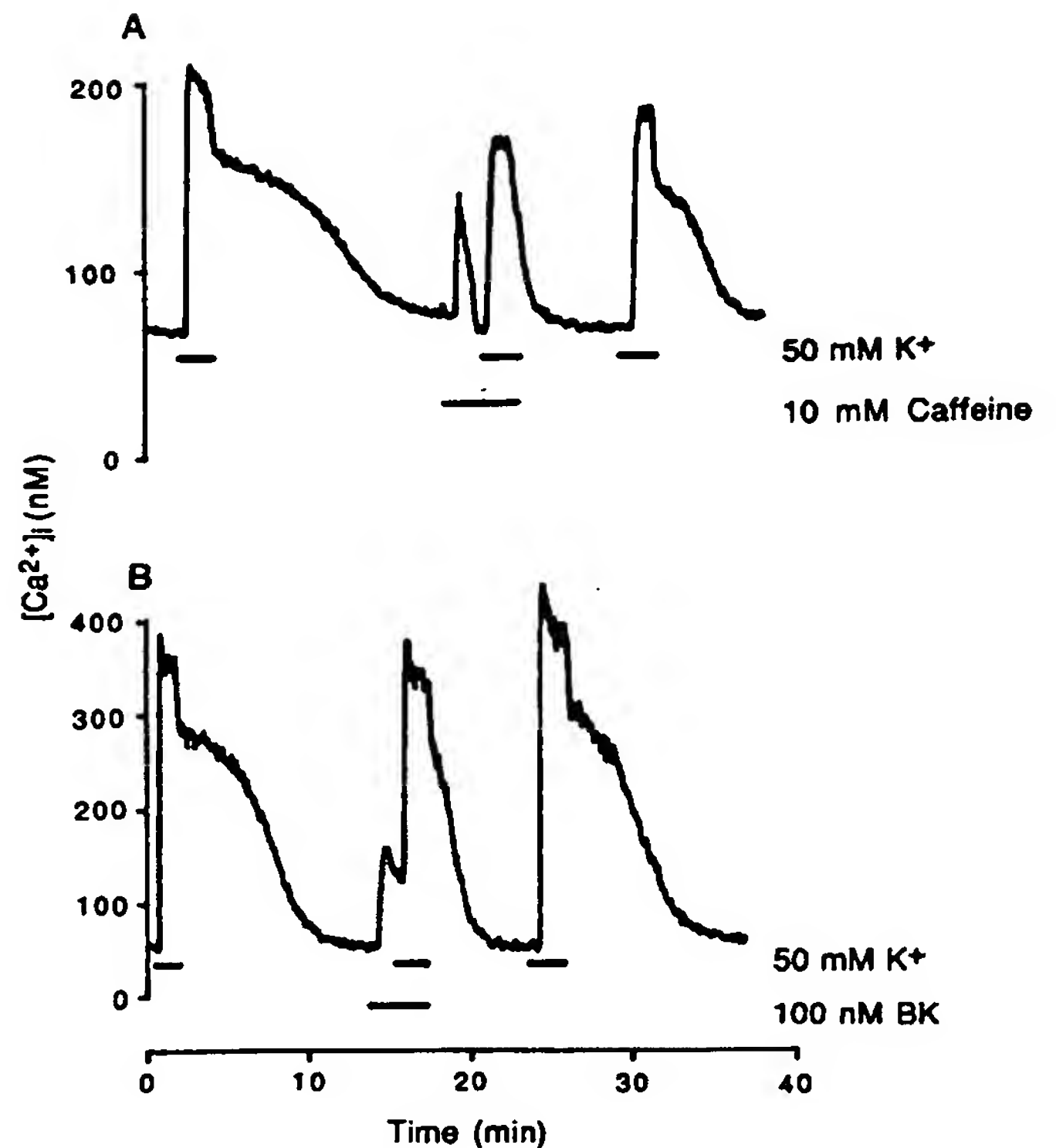




**Figure 10.** Separation of BK and caffeine sensitivity of the intracellular  $\text{Ca}^{2+}$  stores in sensory neurons. A DRG soma (A) and process (B) were perfused with  $\text{Ca}^{2+}$ -free ( $20 \mu\text{M}$  EGTA) solution containing either 100 nM BK or 10 mM caffeine as indicated. Depolarization-induced  $\text{Ca}^{2+}$  influx was produced by perfusion with 50 mM  $\text{K}^{+}$  in normal  $\text{Ca}^{2+}$ -containing medium. The BK response in the process (B) is an especially large  $[\text{Ca}^{2+}]_i$  transient chosen because of the low noise in the record.

second messengers (Hong and Deykin, 1982; Miller, 1987b). Such substances could mediate the known effects of BK on DRG excitability (Miller, 1987b). In the present study we have demonstrated that in DRG cells, as in other cases, BK does indeed stimulate the production of several phospholipid-derived second messengers, suggesting that it activates phospholipases C and  $\text{A}_2$ . As in other cases, the  $\text{IP}_3$  produced probably plays an important role in mediating the BK-induced changes in  $\text{Ca}^{2+}$  metabolism we have observed. The role of arachidonic acid in this respect is unclear, and we have not investigated it further. However, as we shall discuss, DAG and eicosanoids derived from the metabolism of arachidonic acid are clearly essential in producing the overall excitatory effects of BK on DRG cells. This is primarily due to their effects on  $\text{Ca}^{2+}$  and  $\text{K}^{+}$  conductances (see below).

The general pattern of phospholipid metabolism produced by BK in DRG cells is similar to that previously reported in other tissues including a DRG  $\times$  neuroblastoma clonal cell line, F11 (Francel et al., 1987). One difference concerns the effects of PTX. In DRG cells, PTX had no effect on  $\text{IP}_3$  production. However, in F-11 cells a partial inhibitory effect was observed (Francel et al., 1987). Although it is clear that agonist-induced  $\text{IP}_3$  synthesis utilizes a G-protein of some sort, the PTX sensitivity of this



**Figure 11.** Intracellular stores and  $[\text{Ca}^{2+}]_i$  buffering in sensory neurons. Increases in the  $[\text{Ca}^{2+}]_i$  in DRG cell bodies were induced by perfusion with depolarizing (50 mM  $\text{K}^{+}$ ) medium in normal  $\text{Ca}^{2+}$  as indicated by the horizontal bars. The cells were perfused with  $\text{Ca}^{2+}$ -free ( $20 \mu\text{M}$  EGTA) solution during the recovery to basal  $[\text{Ca}^{2+}]_i$  levels. Caffeine, 10 mM (A), or BK, 100 nM (B), was added to the perfusion solution prior to and during the depolarization as indicated by the horizontal bars.

entity (or entities) has been found to vary widely (Miller, 1988). Actually, it is only in a minority of cases that PTX sensitivity of the  $\text{IP}_3$  generating process has been observed (Berridge, 1987; Miller, 1988). In the particular case of BK, both PTX-sensitive and -insensitive  $\text{IP}_3$  production has been reported (Miller, 1988). Whether this means that BK receptors can be coupled to different G-proteins (Burch and Axelrod, 1987) or whether it reflects the known heterogeneity of BK receptors is as yet unclear (Manning et al., 1986).

In many cell types,  $\text{IP}_3$  generation subsequently leads to the mobilization of intracellular  $\text{Ca}^{2+}$  (Berridge, 1987). This also appears to be true in DRG cells. BK clearly produced increases in  $[\text{Ca}^{2+}]_i$  in many cells even in  $\text{Ca}^{2+}$ -free medium. It is interesting to note that responses to BK appeared to be of about the same magnitude even if  $\text{Ca}^{2+}$ -containing medium was used (S. Thayer and R. Miller, unpublished observations). This indicates that BK does not produce extensive  $\text{Ca}^{2+}$  influx. Thus, any subsequent metabolism of  $\text{IP}_3$  to  $\text{IP}_2$  and associated  $\text{Ca}^{2+}$  influx in these neurons may not be of great importance (Higashida and Brown, 1986a; Irvine and Moore, 1986; Morris et al., 1987). It should be pointed out that, although we imagine that the increases in  $[\text{Ca}^{2+}]_i$  we have observed in response to BK involve  $\text{IP}_3$ , the ability of  $\text{IP}_3$  to actually mobilize  $\text{Ca}^{2+}$  in DRG cells (or indeed in any other type of vertebrate neuron) remains to be directly demonstrated (see, however, Freedman and Aghajanian, 1987; Shah et al., 1987). BK did not increase  $[\text{Ca}^{2+}]_i$  in all



neurons studied. This is probably a reflection of some degree of heterogeneity in the cultures. BK receptors probably exist primarily on small polymodal nociceptors and may not be present on the cell bodies or processes of all DRG neurons.

In our studies, it appeared that BK-sensitive  $\text{Ca}^{2+}$  stores existed in both the cell soma and processes. These stores may differ from those that can be mobilized by high concentrations of methylxanthines. It is well known that caffeine can mobilize an intracellular  $\text{Ca}^{2+}$  store in various types of muscle (Carafoli, 1987). It is thought that this type of store can participate in the phenomenon of " $\text{Ca}^{2+}$ -induced  $\text{Ca}^{2+}$  release," which may be a way of amplifying the  $\text{Ca}^{2+}$  signal in these cells (Barcenas-Ruiz and Weir, 1987; Carafoli, 1987). By implication, such stores have also been thought to exist in bullfrog sympathetic neurons (Kuba, 1980), and this fact has now been directly demonstrated using imaging techniques (Lipscombe et al., 1987). Studies we have carried out using rat superior cervical ganglion neurons have also demonstrated the existence of caffeine-sensitive stores which can be blocked by both ryanodine and dantrolene (Thayer et al., 1987b). Presumably, therefore, these stores are similar to those found in muscle. The present studies demonstrate that methylxanthine-sensitive stores also seem to exist in DRG cells where again they can be blocked by ryanodine. It has been observed that in both rat and bullfrog sympathetic neurons caffeine-sensitive stores seem to exist primarily in the cell soma (Lipscombe et al., 1987; Thayer et al., 1987b). This is also true in DRG cells. We also noted that in many instances it was possible to produce BK responses from a cell process that was completely insensitive to caffeine. If we make the assumption that the BK-induced rise in  $[\text{Ca}^{2+}]_i$  results from BK-stimulated  $\text{IP}_3$  synthesis and subsequent mobilization of  $\text{Ca}^{2+}$  stores, then it appears likely that the caffeine and  $\text{IP}_3$ -sensitive stores are separate entities. Indeed, such a separation of  $\text{IP}_3$ - and caffeine-sensitive stores has also been observed in non-neuronal tissues (Kanaide et al., 1987).

It is clear from our data that the manner in which DRG cells buffer  $\text{Ca}^{2+}$  is quite extraordinary. In comparison with all other types of peripheral and central neurons we have studied, the shape of the  $\text{Ca}^{2+}$  transient produced upon depolarization is remarkable. Following a rapid phase of buffering,  $[\text{Ca}^{2+}]_i$  falls very slowly. The tardiness of  $\text{Ca}^{2+}$  buffering in these cells is also observed on a much shorter time scale when voltage-clamp steps are used to depolarize the cells (Thayer et al., 1988). It is interesting to note that Jia and Nelson (1986) predicted that DRG cells might buffer  $\text{Ca}^{2+}$  slowly on the basis of the paucity of mitochondria they observed in DRG nerve terminals. However, it is not clear that this is actually the reason involved. Thus, the  $\text{Ca}^{2+}$  transients observed in cell processes seem considerably faster than those observed in the cell soma. Our results indicate that the rate of  $\text{Ca}^{2+}$  buffering may be critically dependent on the availability of intracellular  $\text{Ca}^{2+}$  storage sites. Thus, if we discharged either the BK- or caffeine-sensitive stores,  $\text{Ca}^{2+}$  buffering in the cell soma became relatively fast. This seems reasonable.  $\text{K}^+$  depolarization, BK, and caffeine all raise  $[\text{Ca}^{2+}]_i$  into the submicromolar range but seldom higher. Thus, a relatively high-affinity  $\text{Ca}^{2+}$  buffer is probably most important for the rapid removal of the accumulated  $\text{Ca}^{2+}$ . Mitochondria and exchange systems are generally of lower affinity than intracellular  $\text{Ca}^{2+}$  storage sites or the cell membrane  $\text{Ca}^{2+}$  ATPase (Baker and DiPollo, 1984; Carafoli, 1987). We presume that after loading during elevated  $\text{K}^+$  depolarization, the  $\text{Ca}^{2+}$  buffering that occurs rapidly mostly reflects  $\text{Ca}^{2+}$  entry into vacant stores. However,

when these are full, the DRG soma appears to be able to cope with the remaining load only very slowly. In contrast, if storage sites have been vacated by using caffeine or BK, then once the  $\text{Ca}^{2+}$  released has been disposed of, presumably by expulsion from the cell, buffering of a depolarization-induced  $\text{Ca}^{2+}$  load can occur more rapidly.

Another observation we have made relates to the degree of increase in  $[\text{Ca}^{2+}]_i$  produced by depolarization of different portions of the DRG neuron. In rat superior cervical ganglion neurons we observed that  $\text{K}^+$  depolarization increased  $[\text{Ca}^{2+}]_i$  to the same extent in all portions of the cell (Thayer et al., 1987a, b). However, this is clearly not the case in DRG neurons. In DRG cells the overall density of  $\text{Ca}^{2+}$  channels in the cell processes is apparently lower than in the cell soma. However, it may well be that high concentrations of  $\text{Ca}^{2+}$  channels exist in certain portions of the DRG processes from which neurotransmitter release actually occurs and that they are scarce in other portions of the process. Such a question is most easily answered using imaging techniques. It is also possible that  $\text{Ca}^{2+}$  buffering in processes is more efficient than in the cell soma.

The increase in  $[\text{Ca}^{2+}]_i$  produced by BK may be important in several respects. Thus, in NG108-15 cells mobilization of  $\text{Ca}^{2+}$  from intracellular stores does lead to neurotransmitter release (Higashida, 1987). Furthermore, as a number of ionic conductances are sensitive to  $\text{Ca}^{2+}$ , BK-induced increases in  $[\text{Ca}^{2+}]_i$  may help to regulate DRG excitability. It should further be noted that BK stimulates the release of DAG and arachidonic acid. The former has been observed to inhibit  $\text{Ca}^{2+}$  currents in DRG cells (Rane and Dunlap), and the latter can be metabolized to various eicosanoids that have been shown to alter  $\text{K}^+$  channels in DRG cells through the production of further second messengers such as cyclic AMP (Weinreich, 1986; Wonderlein and Weinreich, 1986).

## References

- Baccaglini, P. T., and P. G. Hogan (1983) Some rat sensory neurons in culture express characteristics of differentiated pain sensory cells. *Proc. Natl. Acad. Sci. USA* 80: 594–598.
- Baker, P. F., and R. DiPollo (1984) Axonal calcium and magnesium homeostasis. *Curr. Top. Membr. Transport* 22: 195–249.
- Barcenas-Ruiz, L., and W. G. Weir (1987) Voltage-dependence of intracellular  $[\text{Ca}^{2+}]_i$  transients in guinea-pig ventricular myocytes. *Circ. Res.* 61: 148–154.
- Beny, J. L., P. Brunet, and H. Huggel (1987) Interaction of bradykinin and des-Arg<sup>9</sup>-bradykinin with isolated pig coronary arteries: Mechanical and electrophysiological events. *Regul. Pep.*, 17: 181–190.
- Berridge, M. J. (1987) Inositol trisphosphate and diacylglycerol: Two interacting second messengers. *Annu. Rev. Biochem.* 56: 159–195.
- Berridge, M. J., C. P. Downes, and M. R. Hanley (1982) Lithium amplifies agonist-dependent phosphatidylinositol responses in brain and salivary glands. *Biochem. J.* 206: 587–595.
- Billah, M. M., and E. G. Lapetina (1982) Rapid decrease of phosphatidylinositol 4,5-bisphosphate in thrombin-stimulated platelets. *J. Biol. Chem.* 257: 12705–12708.
- Burch, R. M., and J. Axelrod (1987) Dissociation of bradykinin induced prostaglandin formation from phosphatidylinositol turnover in Swiss 3T3 fibroblasts: Evidence for G-protein regulation of phospholipase A<sub>2</sub>. *Proc. Natl. Acad. Sci. USA* 84: 6374–6379.
- Carafoli, E. (1987) Intracellular calcium homeostasis. *Annu. Rev. Biochem.* 56: 395–435.
- Erdos, E. G. (1979) Bradykinin, kallidin and kallikrein. In *Handbook of Experimental Pharmacology*, Vol. 25, Springer-Verlag, Berlin.
- Ewald, D. A., H. J. G. Matthies, T. M. Perney, M. W. Walker, and R. J. Miller (1988) The effect of down regulation of protein kinase C on the inhibitory modulation of dorsal root ganglion neuron  $\text{Ca}^{2+}$  currents by neuropeptide Y. *J. Neurosci.* 8: 2447–2451.
- Fabiato, A., and F. Fabiato (1979) Calculator programs for computing the composition of the solutions containing multiple metals and li-

- gands used for experiments in skinned muscle cells. *J. Physiol. (Paris)* 75: 463–505.
- Fowler, J. C., R. Greene, and D. Weinreich (1985) Two calcium sensitive spike afterhyperpolarizations in visceral sensory neurones of the rabbit. *J. Physiol. (Lond.)* 365: 59–75.
- Francel, P. C., and G. Dawson (1986) Bradykinin induces a rapid release of inositol trisphosphate from a neuroblastoma hybrid cell line NCB-20 that is not antagonized by enkephalin. *Biochem. Biophys. Res. Commun.* 135: 507–514.
- Francel, P. C., R. J. Miller, and G. Dawson (1987) Modulation of bradykinin induced inositol trisphosphate release in a novel neuroblastoma × dorsal root ganglion sensory neuron cell line (F-11). *J. Neurochem.* 48: 1632–1639.
- Freedman, J. E., and G. K. Aghajanian (1987) Role of phosphoinositide metabolites in the prolongation of afterhyperpolarizations by  $\alpha_1$ -adrenoceptors in rat dorsal raphe neurons. *J. Neurosci.* 7: 3897–3906.
- Griendling, K. K., S. E. Rittenshouse, T. A. Brock, L. S. Ekstein, M. A. Gimbrone, Jr., and R. W. Alexander (1986) Sustained diacylglycerol formation from inositol phospholipids in angiotensin II-stimulated vascular smooth muscle cells. *J. Biol. Chem.* 261: 5901–5906.
- Grynkiewicz, G., M. Poenie, and R. Y. Tsien (1985) A new generation of  $\text{Ca}^{2+}$  indicators with greatly improved fluorescence properties. *J. Biol. Chem.* 260: 3440–3450.
- Higashi, H., N. Ueda, S. Nishi, J. P. Gallagher, and P. Shinnick-Gallagher (1982) Chemoreceptors for serotonin (5-HT), acetylcholine (ACh), bradykinin (BK), histamine (H) and  $\gamma$ -aminobutyric acid (GABA) on rabbit visceral afferent neurons. *Brain Res. Bull.* 8: 23–32.
- Higashida, H. (1987) Presynaptic role of inositol-1,4-trisphosphate and diacylglycerol in bradykinin induced acetylcholine release at NG108-15 neuroblastoma hybrid myotube synapses. *Soc. Neurosci. Abstr.* 13: 67.
- Higashida, H., and D. A. Brown (1986a) Membrane current responses to intracellular injections to inositol-1,3,4,5-tetrakisphosphate and inositol-1,3,4-trisphosphate in NG108-15 hybrid cells. *FEBS Lett.* 208: 283–286.
- Higashida, H., and D. A. Brown (1986b) Two polyphosphatidylinositol metabolites control two  $\text{K}^+$ -currents in a neuronal cell. *Nature* 323: 333–335.
- Higashida, H., and D. A. Brown (1987) Bradykinin inhibits potassium ( $\text{M}$ ) currents in N1E-115 neuroblastoma cells. *FEBS Lett.* 220: 302–306.
- Higashida, H., R. A. Streaty, W. Klee, and M. Nirenberg (1986) Bradykinin-activated transmembrane signals are coupled via  $\text{N}_\text{o}$  or  $\text{N}_\text{i}$  to production of inositol-1,4,5-trisphosphate, a second messenger in NG108-15 neuroblastoma-glioma hybrid cells. *Proc. Natl. Acad. Sci. USA* 83: 942–947.
- Hong, S. L., and D. Deykin (1982) Activation of phospholipases  $\text{A}_2$  and C in pig aortic endothelial cells synthesizing prostacyclin. *J. Biol. Chem.* 257: 7151–7154.
- Irvine, R. F., and R. M. Moore (1986) Microinjection of inositol 1,3,4,5-tetrakisphosphate activates sea urchin eggs by a mechanism dependent on external  $\text{Ca}^{2+}$ . *Biochem. J.* 240: 917–920.
- Jackson, T. R., T. J. Hallam, C. P. Downes, and M. R. Hanley (1987) Receptor coupled events in bradykinin action: Rapid production of inositol phosphates and the regulation of cytosolic free  $\text{Ca}^{2+}$  in a neural cell line. *EMBO J.* 6: 49–54.
- Jia, M., and P. G. Nelson (1986) Calcium currents and transmitter output in cultured spinal and dorsal root ganglion neurones. *J. Neurophysiol.* 56: 1242–1256.
- Kanaide, H., Y. Shogakiuchi, and M. Nakamura (1987) The norepinephrine-sensitive  $\text{Ca}^{2+}$ -storage site differs from the caffeine-sensitive site in vascular smooth muscle of the rat aorta. *FEBS Lett.* 214: 130–134.
- Kuba, K. (1980) Release of calcium ions linked to the activation of a potassium conductance in a caffeine treated sympathetic neurone. *J. Physiol. (Lond.)* 298: 251–269.
- Lipscombe, D., D. V. Madison, M. Poenie, H. Reuter, R. Y. Tsien, and R. W. Tsien (1987) Spatial distribution of calcium channels and cytosolic calcium transients in growth cones and cell bodies of sympathetic neurones. *Proc. Natl. Acad. Sci. USA* 85: 2398–2402.
- Manning, D. C., and S. H. Snyder (1983) [ $^3\text{H}$ ]-bradykinin in receptor localization in spinal cord and sensory ganglia, evidence for a role in primary afferent function. *Soc. Neurosci. Abstr.* 9: 590.
- Manning, D. C., S. H. Snyder, J. F. Kachur, R. J. Miller, and M. Field (1982) Bradykinin receptor mediated chloride secretion in intestinal function. *Nature* 299: 256–259.
- Manning, D. C., R. Vavrek, J. M. Stewart, and S. H. Snyder (1986) Two bradykinin binding sites with picomolar affinities. *J. Pharmacol. Exp. Ther.* 237: 504–512.
- Matthies, H., H. C. Palfrey, L. D. Hirning, and R. J. Miller (1987) Down regulation of protein kinase C in neuronal cells: Effects on neurotransmitter release. *J. Neurosci.* 7: 1198–1206.
- Miller, R. J. (1987a) Multiple calcium channels and neuronal function. *Science* 235: 46–52.
- Miller, R. J. (1987b) Bradykinin highlights the role of phospholipid metabolism in the control of nerve excitability. *Trends Neurosci.* 10: 226–228.
- Miller, R. J. (1988) G-proteins flex their muscles. *Trends Neurosci.* 11: 3–6.
- Morris, A. P., D. V. Gallacher, R. F. Irvine, and O. H. Petersen (1987) Synergism of inositol trisphosphate and tetrakisphosphate in activating  $\text{Ca}^{2+}$  dependent  $\text{K}^+$  channels. *Nature* 330: 653–655.
- Murphy, S. M., S. A. Thayer, and R. J. Miller (1987) Effects of excitatory amino acids on  $[\text{Ca}^{2+}]$ , on single striatal neurons. *J. Neurosci.* 7: 4145–4158.
- Neering, I. R., and R. W. McBurney (1984) Role for microsomal  $\text{Ca}$  storage in mammalian neurones. *Nature* 309: 158–160.
- Osugi, T., T. Imaizumi, A. Mizushima, S. Uchida, and Y. Yoshida (1986a) 1-oleoyl-2-acetyl-glycerol and phorbol diester stimulate  $\text{Ca}^{2+}$  influx through  $\text{Ca}^{2+}$  channels in neuroblastoma × glioma hybrid NG108-15 cells. *Eur. J. Pharmacol.* 126: 47–51.
- Osugi, T., S. Uchida, T. Imaizumi, and H. Yoshida (1986b) Bradykinin induced intracellular  $\text{Ca}^{2+}$  elevation in neuroblastoma × glioma hybrid NG108-15 cells: Relationship to the action of inositol phospholipid metabolites. *Brain Res.* 379: 84–89.
- Perney, T. M., L. D. Hirning, S. E. Leeman, and R. J. Miller (1986) Multiple  $\text{Ca}^{2+}$  channels mediate neurotransmitter release from peripheral neurones. *Proc. Natl. Acad. Sci. USA* 83: 6656–6659.
- Rane, S. G., and K. Dunlap (1986) Kinase C activator 1,2-oleoyl-acetyl-glycerol attenuates voltage dependent calcium current in sensory neurones. *Proc. Natl. Acad. Sci. USA* 83: 184–188.
- Rang, H., and J. M. Ritchie (1987) Activation of protein kinase C causes a depolarization of the rat vagus nerve associated with increased sodium conductance. *J. Physiol. (Lond.)* 391: 789.
- Reiser, G., and B. Hamprecht (1985) Bradykinin causes a transient rise of intracellular  $\text{Ca}^{2+}$  activity in cultured neuronal cells. *Pfluegers Arch.* 405: 260–264.
- Shah, J., R. S. Cohen, and H. C. Pant (1987) Inositol trisphosphate-induced calcium release in brain microsomes. *Brain Res.* 419: 1–6.
- Thayer, S. A., L. D. Hirning, and R. J. Miller (1987a) The distribution of multiple types of  $\text{Ca}^{2+}$  channels in rat sympathetic neurones *in vitro*. *Mol. Pharmacol.* 32: 579–586.
- Thayer, S. A., L. D. Hirning, K. M. Harris, and R. J. Miller (1987b) Distribution of multiple  $\text{Ca}^{2+}$  channel types and intracellular  $\text{Ca}^{2+}$  stores in single central and peripheral neurons. *Soc. Neurosci. Abstr.* 13: 1010.
- Thayer, S. A., M. Sturek, and R. J. Miller (1988) Measurement of neuronal  $\text{Ca}^{2+}$  transients using simultaneous microfluorimetry and electrophysiology. *Pfluegers Arch.* 412: 216–223.
- Van Calcar, D., and R. Heumann (1987) Nerve growth factor potentiates the agonist stimulated accumulation of inositol phosphates in PC-12 pheochromocytoma cells. *Eur. J. Pharmacol.* 135: 259–260.
- Weinreich, D. (1986) Bradykinin inhibits a slow spike afterhyperpolarization in visceral sensory afferents. *Eur. J. Pharmacol.* 132: 61–63.
- Wonderlein, W. F., and D. Weinreich (1986)  $\text{Ca}^{2+}$  dependent outward currents activated by  $\text{Ca}^{2+}$  injection in visceral sensory neurones. *Soc. Neurosci. Abstr.* 12: 1200.
- Yaksh, T. L., and D. L. Hammond (1982) Peripheral and central substrates involved in the transmission of nociceptive information. *Pain* 13: 1–46.
- Yano, K., H. Higashida, R. Inoue, and Y. Nozawa (1984) Bradykinin induced rapid breakdown of phosphatidylinositol 4,5-bisphosphate in neuroblastoma × glioma hybrid NG108-15 cells. *J. Biol. Chem.* 259: 10201–10207.



# Purification and Characterization of an Abnormal Factor IX (Christmas Factor) Molecule

## FACTOR IX CHAPEL HILL

KUO-SAN CHUNG, DEAN A. MADAR, JONATHAN C. GOLDSMITH, HENRY S. KINGDON, and HAROLD R. ROBERTS, *Departments of Medicine, Pathology, Biochemistry, and Chemistry, The University of North Carolina, Chapel Hill, North Carolina 27514*

**ABSTRACT** Human Factor IX (Christmas factor) was isolated from the plasma of a patient with mild hemophilia B. The patient's plasma contained 5% Factor IX clotting activity but 100% Factor IX antigenic activity as determined by immunological assays, which included inhibitor neutralization and a radioimmunoassay for Factor IX. This abnormal Factor IX is called Factor IX Chapel Hill (Factor IX<sub>CH</sub>). Both normal Factor IX and Factor IX<sub>CH</sub> have tyrosine as the NH<sub>2</sub>-terminal amino acid. The two proteins have a similar molecular weight, a similar amino acid analysis, the same number of gamma-carboxyglutamic acid residues (10  $\gamma$ -carboxyglutamic acid residues), and a similar carbohydrate content. Both exist as a single-chain glycoprotein in plasma. The major difference between normal Factor IX and Factor IX<sub>CH</sub> is that the latter exhibits delayed activation to Factor IXa in the presence of Factor XIa and Ca<sup>2+</sup>. Thus, Factor IX<sub>CH</sub> differs from other previously described abnormal Factor IX molecules.

## INTRODUCTION

Factor IX (plasma thromboplastin component, Christmas factor) is a clotting factor which participates in the middle phase of intrinsic blood coagulation. Diminished activity of this factor results in hemophilia B. Factor IX is a glycoprotein that has been isolated from both the human and bovine species and many of the characteristics of the normal factor have been published (1-8). In the circulation, Factor

IX exists as a single-chain zymogen which is converted to an active serine protease in the presence of Factor XIa (plasma thromboplastin antecedent) and Ca<sup>2+</sup> ions (9). During activation, the Factor IX zymogen undergoes proteolysis in a two-stage reaction that results in a two-chain molecule composed of a light and heavy chain linked by disulfide bonds. The first step in the proteolysis is the cleavage of a specific arginine-alanine peptide bond which results in an inactive intermediate product. The second cleavage splits an arginine-valine peptide bond with the release of a carbohydrate-rich activation peptide and an active Factor IX enzyme (10, 11). The active site serine resides on the heavy chain (11).

Activated Factor IX (Factor IXa), in the presence of phospholipids, thrombin-modified Factor VIII, and calcium ions, converts Factor X to Factor Xa (2). However, the precise molecular interactions of Factor IXa with calcium, phospholipid, Factor VIII, and Factor X are not completely understood. Elucidation of the nature of the defect in abnormal Factor IX molecules found in some patients with hemophilia B (Christmas disease) may provide insights into function of normal Factor IX in hemostasis and its interaction with other clotting factors. Roberts et al. (12), and other investigators (13-19) have demonstrated that some patients with hemophilia B possess an abnormal Factor IX molecule which has low to undetectable clotting activity, but which is present in relatively normal amounts when measured by immunological techniques (12, 20-22). Because they contain cross-reacting material (CRM)<sup>1</sup> to specific homologous and heterologous

An abstract of this work appeared in 1976. *Blood*. 48: 974. (Abstr.)

Dr. Goldsmith was supported by a National Hemophilia Foundation Fellowship.

Received for publication 3 February 1978 and in revised form 13 July 1978.

<sup>1</sup> Abbreviations used in this paper: CRM, cross-reacting material; CRM<sup>+</sup>, CRM positive variants; CRM<sup>-</sup>, CRM negative variants; CRM<sup>R</sup>, CRM reduced variants; Factor IX<sub>CH</sub>, Factor IX Chapel Hill; SDS, sodium dodecyl sulfate.



anti-Factor IX antibodies, these patients have been termed CRM<sup>+</sup> (CRM positive variants). They can be distinguished from other hemophilia B patients who have reduced, but not absent, levels of both Factor IX antigen and Factor IX clotting activity (CRM<sup>R</sup>; CRM reduced variants). The CRM<sup>+</sup> patients can also be distinguished from those hemophilia B patients in whom Factor IX antigen and Factor IX clotting activity are absent or undetectable (CRM<sup>-</sup>; CRM negative variants) (22).

The purpose of this paper is to describe the isolation and characteristics of an abnormal Factor IX molecule from a mildly affected CRM<sup>+</sup> hemophilia B patient. This abnormal Factor IX has been named Factor IX Chapel Hill (Factor IX<sub>CH</sub>) (23). Factor IX<sub>CH</sub> is compared with purified normal human Factor IX in terms of its molecular weight, NH<sub>2</sub>-terminal amino acid, content of  $\gamma$ -carboxyglutamic acid residues, carbohydrate content, and activation by Factor XIa and Ca<sup>2+</sup>. The major abnormality in Factor IX<sub>CH</sub> appears to be defective activation by Factor XIa and Ca<sup>2+</sup>, although abnormalities in interactions with other clotting factors cannot be excluded.

## METHODS

**Materials.** The patient with Factor IX<sub>CH</sub> is a 30-yr-old white male who has been followed at the University of North Carolina Hospital for 16 yr. He has a mild bleeding tendency and a Factor IX level of 5% of normal when tested in one-stage clotting assays. With a specific antibody to Factor IX, however, the patient was found to have 100% of Factor IX when measured by an antibody neutralization test and a radioimmunoassay for Factor IX (*vide infra*). All other clotting factors were normal. Plasma for purification of Factor IX<sub>CH</sub> was obtained from this patient who was previously shown to be CRM<sup>+</sup> (24). The patient was plasmapheresed every 2 wk to 1 mo, after informed consent, until sufficient plasma was obtained for the studies outlined below. The blood was handled in a manner similar to that for normal blood as described below. The resultant plasma was stored at -70°C and thawed at 37°C just before use. Normal, pooled human plasma was obtained from the Blood Bank of the North Carolina Memorial Hospital, Chapel Hill, N. C. Human blood was collected in plastic bags that contained standard citrate phosphate dextrose anticoagulant and centrifuged at 3,000 g for 30 min at 4°C to render the plasma platelet poor and free of erythrocytes. The plasma was then recentrifuged under the same conditions. Plasmas used for the purification of normal Factor IX were either fresh or stored for <2 mo at -70°C.

Alumina C- $\gamma$  gel was purchased from Calbiochem, San Diego, Calif. Benzamidinium hydrochloride and cyanogen bromide were purchased from Eastman Organic Chemicals, Div. Eastman Kodak Co., Rochester, N. Y. Heparin sodium salt, bovine serum albumin, dithioerythritol, dansyl chloride, dansyl amino acids, galactose, mannose, glucose, fucose, N-acetylneuraminic acid, and amino sugars were obtained from Sigma Chemical Co., St. Louis, Mo. DEAE-cellulose was Whatman DE-52 (Whatman Chemicals, Div. W & R Balston, Maidstone, Kent, England). DEAE-Sephadex (A-50), Sephadex G-150, Ribonuclease A, chymotrypsinogen A, and ovalbumin were purchased from Pharmacia Fine Chemicals, Uppsala,

Sweden. Bovine brain cephalin (Thromboplastin) was purchased from Ortho Pharmaceuticals Corporation, Raritan, N. J. Thromboplastin reagent (Simplastin) was purchased from General Diagnostics, Warner Lambert Company, Morris Plains, N. J. Agarose (Bio-Gel A-15 m, 100-200 mesh), Aminex A-4 Resin,  $\beta$ -mercaptoethanol, and materials used in gel electrophoresis were obtained from Bio-Rad Laboratories, Richmond, Calif. Celite 545 was purchased from Johns-Manville, Denver, Colo. All other chemicals were of reagent grade or better.

**Protein concentrations.** The Lowry method (25), with bovine serum albumin as a standard, was used for the estimation of protein concentration for the early stages of purification of Factor IX. Protein concentration on the final product was estimated by optical adsorption at 280 nm with experimentally determined extinction coefficients (*vide infra*).

**Coagulation assays.** 1 U of Factor VII, VIII, IX, or X clotting activity was defined as that amount in 1 ml of normal, pooled human plasma. A one-stage assay, based on the partial thromboplastin time, was used for the measurement of Factors VIII and IX (26, 27). Factors X and VII were measured by a one-stage assay based on the prothrombin time (28). Prothrombin was assayed with a two-stage method (29). Factor XI was assayed with artificially depleted human plasma (30) as well as plasma congenitally deficient in Factor XI (31). Factor IX immunological activity was assayed by an inhibitor-neutralization technique (12, 21). Factor IX was also measured by a radioimmunoassay with purified normal human Factor IX and a human anti-Factor IX antibody (32).

**Preparation of heparin-agarose.** Heparin-agarose was prepared by a modification of the method described by Porath (33) and Cuatrecasas (34).

**Electrophoresis.** Disc gel electrophoresis was performed in 0.2 M N,N-bis[2-hydroxyethyl]glycine-imidazole buffer at pH 7.8 for nondenatured protein and 0.025 M Tris-glycine buffer at pH 8.3 for sodium dodecyl sulphate-treated protein (35). 7.5 and 10% polyacrylamide gels were used for analytical separation of the proteins with 3% polyacrylamide as the stacking gel.  $\beta$ -Mercaptoethanol and dithioerythritol were used as reducing agents. Immuno-electrophoresis was performed on 1% agarose gel in 0.075 M Tris-barbital buffer at pH 8.6 which contained 2 mM of calcium lactate.

**Purification of Factor IX<sub>CH</sub>.** 60 ml of cold alumina gel was stirred into 600 ml of cold platelet-poor plasma, obtained from the patient with Factor IX<sub>CH</sub>, for 5 min at 25°C. The plasma was centrifuged and the supernate decanted. The adsorbate was washed twice with cold, normal saline and then eluted with 30 ml 0.3 M (NH<sub>4</sub>)<sub>2</sub>HPO<sub>4</sub> buffer at pH 8.5. The elution was repeated at least once. Pooled eluate was dialyzed against 4 liters of 0.01 M sodium phosphate buffer at pH 7.0. The dialysate was then applied to a DEAE-cellulose (DE-52, Whatman Chemicals) column (2.5  $\times$  25 cm) which was washed extensively with sodium phosphate buffer containing 0.13 M NaCl until the OD (280 nm) of the effluent was <0.01. The column was eluted with a linear gradient of NaCl (from 0.13 M to 0.30 M) with an automatic level sensor. Factor IX<sub>CH</sub>-rich fractions were pooled and dialyzed against 4 liters of 0.05 M sodium citrate at pH 7.0. The dialysate was applied to a DEAE-Sephadex column (2.3  $\times$  12 cm) (Pharmacia Fine Chemicals) which was washed with 0.08 M sodium citrate (pH 7.0) until the OD (280 nm) was <0.01. A linear gradient of sodium citrate from 0.08 to 0.30 M (pH 7.0) was then used to elute the column. Factor IX<sub>CH</sub>-rich fractions were pooled and dialyzed against 4 liters of 0.05 M imidazole-HCl buffer (pH 6.0) overnight, and the dialysate was brought to 2.5 mM CaCl<sub>2</sub> concentration with a 1-M CaCl<sub>2</sub> stock solution. A heparin-agarose column (1.1  $\times$  4.5 cm) was used for the last step of purification. After the Factor IX solution was applied, the column was washed with 250 ml of imidazole-HCl buffer which contained

0.25 M NaCl and 2.5 mM CaCl<sub>2</sub>. Factor IX was eluted with 0.4 M NaCl and 2.5 mM CaCl<sub>2</sub> in the same buffer. The procedures were carried out at 4°C except when stated otherwise. All solutions used in the purification procedure contained 1 mM benzamidine hydrochloride. Purification was monitored by a two-stage Factor IX neutralization assay as well as clotting assays for Factors II, VII, VIII, IX, and X.

**Purification of normal human Factor IX.** Purification of normal Factor IX was similar to that described for Factor IX<sub>CH</sub>.

**Preparation of Factor XIa.** 10 ml of normal, pooled plasma was adsorbed twice with alumina C-γ gel for 30 min at 25°C. After centrifugation, the supernatant plasma was incubated with 200 mg of prewashed Celite 545 (Johns-Manville) in a plastic tube at 37°C for 10 min. The tube was inverted every 1–2 min. The Celite suspension was collected by centrifugation and then washed 3 times with 10 ml of distilled water. Factor XIa was eluted from the Celite with 10 ml of 10% saline, and dialyzed against buffered citrated saline (pH 7.2). Factor XIa was also purified by the method of Lundblad and Kingdon (36) and used for the activation of purified Factor IX.

**Molecular weight.** The molecular weight was estimated by the electrophoretic mobility method with sodium dodecyl sulfate (SDS) disc gel electrophoresis as described above (35).

**Inhibitor neutralization assay and radioimmunoassay for Factor IX.** The inhibitor neutralization assay was performed according to the method of Roberts et al. (12). A radioimmunoassay was carried out according to the method described by Reisner et al. (32).

**Activation of Factor IX by Factor XIa and Ca<sup>2+</sup>.** 3 mg of purified, normal human Factor IX were incubated with 100 μg of partially purified Factor XIa at 37°C in 0.05 M imidazole buffer (pH 6.0) in a final volume of 3 ml which contained 5 mM of CaCl<sub>2</sub>. At selected time intervals an aliquot of the incubation mixture was added to platelet-poor, unactivated Factor IX-deficient substrate plasma obtained from a severely affected CRM<sup>-</sup> hemophilia B patient. A nonactivated, partial thromboplastin time and a Factor IX assay were performed on the final mixture. At the same time intervals, aliquots from the first incubation mixture were evaluated on SDS gel electrophoresis to determine the time of appearance of lower molecular weight components. Factor IX<sub>CH</sub> was also activated in the same manner, and monitored in a similar fashion. Appropriate controls were run by substituting imidazole buffer for Factor XIa, Factor IX, and Ca<sup>2+</sup> in the incubation mixture.

**NH<sub>2</sub>-terminal amino acid analysis.** The NH<sub>2</sub>-terminal amino acid was determined by Edman degradation (37) with a Beckman model 890-C Sequencer (Beckman Instruments, Inc., Fullerton, Calif.). Phenyl isothiocyanate was coupled to the free amino terminal group of the protein by incubation at 55°C for 15 min. Anhydrous heptafluorobutyric acid was used to release the phenyl isothiocyanate derivative of the NH<sub>2</sub>-terminal amino acid which was identified by gas chromatography. The NH<sub>2</sub>-terminal amino acid was also determined by thin layer chromatography with dansyl chloride (38).

**Amino acid composition of human Factor IX.** Amino acid analysis was performed by the method of Spackman et al. (39). Samples in 6 N HCl were sealed in Pyrex (Corning Glass Works, Corning, N. Y.) tubes (Research & Development Products, Berkeley, Calif.) at reduced pressure and hydrolyzed at 110°C for 24 and 72 h. γ-Carboxyglutamic acid was quantitated from alkaline hydrolysates of protein with a modification of the method of Hauschka et al. (40). 1 mg of each of these samples was dissolved in 2 ml of 2 N KOH and sealed at reduced pressure in glass tubes lined with polypropylene. Hydrolysis was carried out at 110°C for 24 h. After alkaline hydrolysis, samples were brought to pH 7.0 by drop-wise addition of 35% HClO<sub>4</sub>. The solutions were evaporated and the

amino acids were eluted from the precipitate cake with aliquots of 0.2 N sodium citrate buffer, pH 2.2. Amino acid analyses were performed on either a modified Phoenix model VG6000-B (Phoenix Precision Instrument, Div. Virtis Co., Inc., Gardiner, N. Y.) or a Beckman model 116 analyzer (Beckman Instruments, Inc.). The determination of cystine as cysteic acid was performed after Moore (41) with bovine prothrombin as a standard to determine the yield of the reaction. The relative amount of tryptophan to tyrosine in Factor IX was determined by a spectrophotometric method (42) in 0.1 N NaOH with a Cary 118 spectrophotometer (Cary Instruments, Fairfield, N. J.). Tryptophan residues per molecule of Factor IX were calculated from the tyrosine residues obtained in the acid hydrolysis.

**Carbohydrate composition.** Neutral sugars were determined by quantitative gas liquid chromatography (43) after being released from the protein by hydrolysis in 2 N H<sub>2</sub>SO<sub>4</sub> at 100°C for 5 h and converted to alditol acetates (44). A colorimetric method with the phenolsulfuric acid reaction and a 1:1 mixture of galactose and mannose as standard was also employed to confirm the result of total hexoses (45). Total hexosamine was determined according to Gardell (46). Sialic acid was assayed with 2-thiobarbituric acid (47).

**Extinction coefficient.** Purified normal Factor IX and Factor IX<sub>CH</sub> were rendered salt-free and lyophilized. The lyophilized proteins were then dissolved in 0.05 M imidazole-HCl buffer (pH 6.0) and the optical absorption at 280 nm was measured and extinction coefficients were calculated.

## RESULTS

Summaries of the purification results of Factor IX<sub>CH</sub> and normal Factor IX are shown in Tables I and II, respectively. Both normal Factor IX and Factor IX<sub>CH</sub> were purified from 17,000- to 18,000-fold with a yield ranging from 47 to 53%.

When the final products were electrophoresed on SDS gels only one band was seen for normal Factor IX and Factor IX<sub>CH</sub> as shown in Fig. 1. The apparent molecular weights according to this technique were 70,000 for both the normal product and Factor IX<sub>CH</sub>. When both normal Factor IX and IX<sub>CH</sub> were run in SDS gels in the presence of β-mercaptoethanol or dithioerythritol, only one band was seen, which in-

TABLE I  
Purification of Factor IX<sub>CH</sub>

Stage	Volume	Total protein	Clotting activity	Yield	Purification
	ml	mg	U	%	-fold
Plasma	600	3.1 × 10 <sup>4</sup>	28.2	100	1
Al (OH) <sub>3</sub>	60	155	26.8	95	190
DE-52	150	5.89	19.5	69	3,640
DEAE-Sephadex	60	2.65	17.7	63	7,340
Heparin-agarose	35	0.96	16.3	58 (54)*	18,700 (17,300)*

\* Antigenic determination in parentheses.

TABLE II  
Purification of Normal Human Factor IX

Stage	Volume	Total protein	Clotting activity	Yield	Purification
	ml	mg	U	%	-fold
Plasma	2,000	$15.3 \times 10^4$	$1.7 \times 10^3$	100	1
Al(OH) <sub>3</sub>	203	$7.3 \times 10^3$	$1.66 \times 10^3$	97.2	204
DEAE-Sephadex	90	$1.1 \times 10^3$	$1.6 \times 10^3$	94.2	1,301
DE-52	180	10.3	$1.0 \times 10^3$	59.0	8,769
Heparin-agarose	45	4.2	$0.8 \times 10^3$	47.0	17,163

indicates that the zymogen form of both molecules is composed of a single polypeptide chain (Fig. 1). To show that the bands on SDS gels were indeed Factor IX, the purified normal Factor IX and Factor IX<sub>CH</sub> were electrophoresed on non-denatured polyacrylamide gels. Two gels for each product were used. One gel was stained with Coomassie Blue (Bio-Rad Laboratories, Richmond, Calif.) to identify the position of the protein band while the other gel was sliced into 1–2 mm lengths and eluted with buffered citrated saline as shown in Figs. 2 and 3. As can be seen in Fig. 2, normal Factor IX clotting activity was eluted in the same position as the protein band. In Fig. 3 both clotting and immunologic assays were used to detect Factor IX<sub>CH</sub>. On immunoelectrophoresis and immunodif-



FIGURE 1 SDS gel electrophoresis of: (1) nonreduced Factor IX<sub>CH</sub>; (2) nonreduced normal Factor IX; (3) reduced Factor IX<sub>CH</sub>; (4) reduced normal Factor IX.

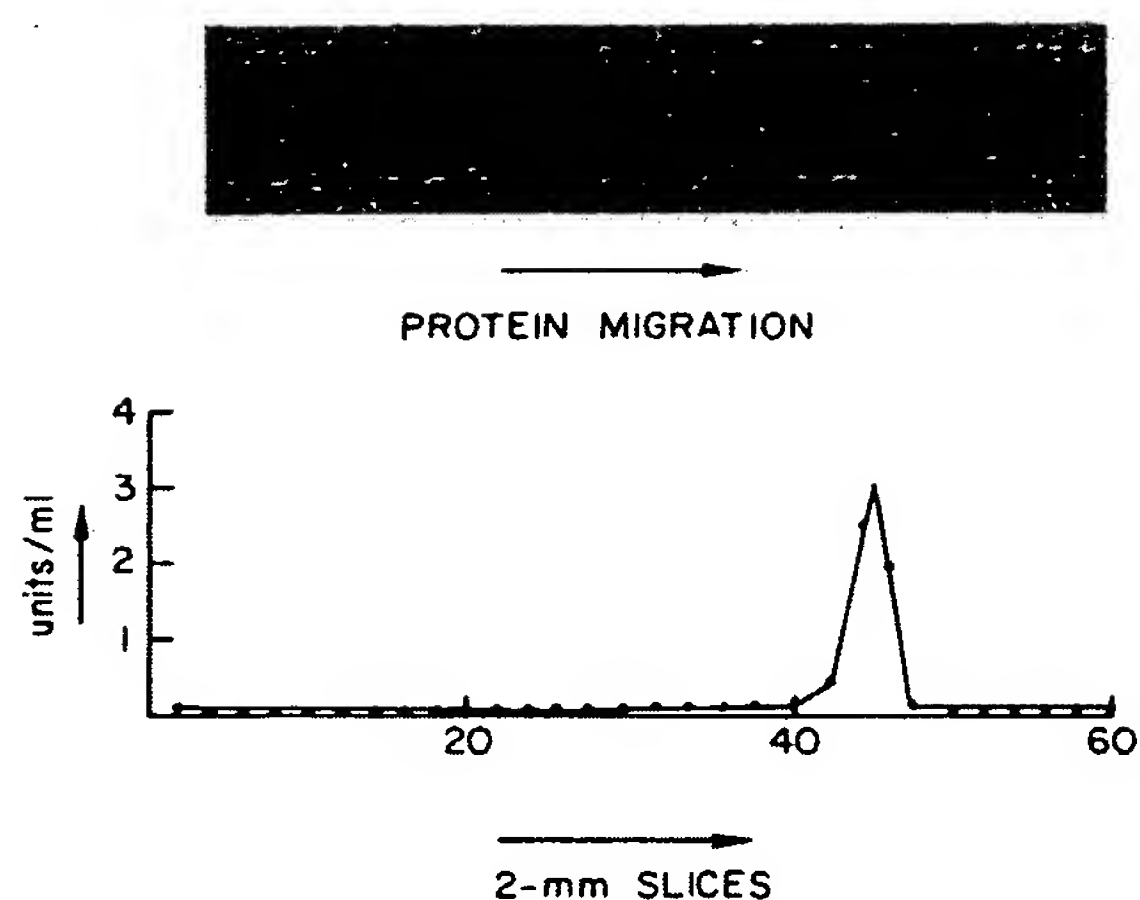


FIGURE 2 Disc gel electrophoresis of normal human Factor IX. 15  $\mu$ g of purified normal human Factor IX was electrophoresed in 7.5% polyacrylamide as shown on the top panel. Protein was identified by staining with Coomassie Blue. The migration was toward the anode. A duplicate gel was sliced into 2-mm segments, as shown in the lower panel, and each segment was crushed and incubated with 0.2 ml 0.15 M NaCl in 0.05 M imidazole-HCl buffer (pH 6.0) at 4°C overnight. The sample was centrifuged and the supernate assayed for Factor IX.

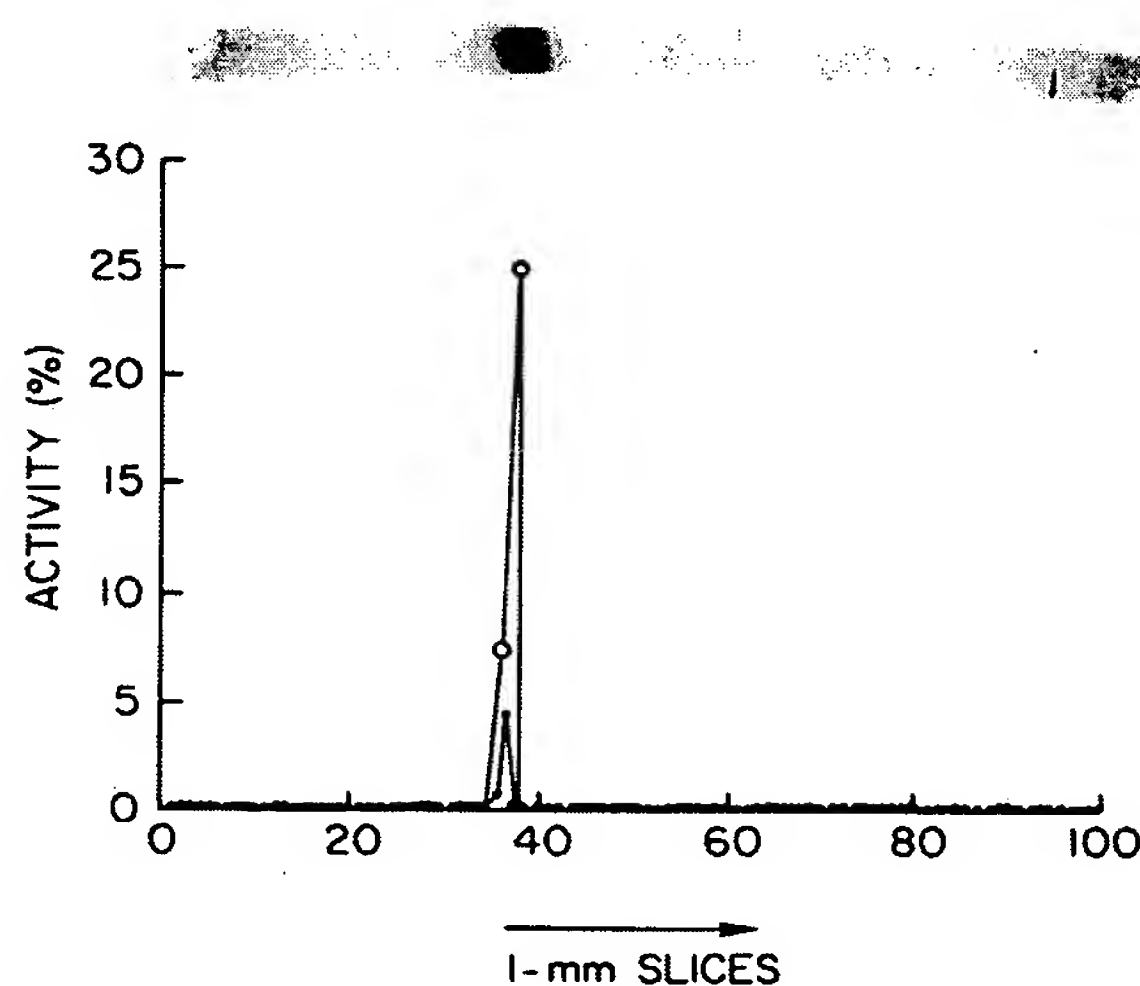


FIGURE 3 Disc gel electrophoresis of Factor IX<sub>CH</sub>. 15  $\mu$ g of purified Factor IX<sub>CH</sub> was electrophoresed on 7.5% polyacrylamide as shown in the top panel. The protein was identified by staining with Coomassie Blue. The migration was toward the anode. A duplicate gel was sliced into 1-mm segments and each segment was crushed and incubated with 0.1 ml 0.15 M NaCl in 0.05 M imidazole buffer (pH 6.0) at 4°C overnight. The Factor IX clotting activity was assayed with the one-stage method and the antigenic activity was assayed by the techniques described in Methods: (○) antigenic activity; (●) clotting activity.



fusion, purified normal Factor IX and IX<sub>CH</sub> were not recognized by antibodies to whole human serum, albumin, prothrombin, and Factor X but were reactive with anti-Factor IX antibodies.

Amino acid analyses of both the normal Factor IX and Factor IX<sub>CH</sub> are shown in Table III. The amino acid compositions of both Factor IX proteins are similar. Both normal Factor IX and Factor IX<sub>CH</sub> contain 10  $\gamma$ -carboxyglutamic acid residues per molecule of Factor IX based on our observed molecular weight for the peptide. They have been subtracted from the glutamic acid content as reported in Table III.

The NH<sub>2</sub>-terminal amino acid of both the normal and abnormal molecule is tyrosine, as determined by both automatic and manual methods on several determinations.

The carbohydrate content of normal Factor IX and IX<sub>CH</sub> is also shown in Table III. The carbohydrate content of each protein is nearly identical except for a

TABLE III  
Amino Acid and Carbohydrate Compositions of Normal Factor IX and Factor IX<sub>CH</sub>

Components	Factor IX	Factor IX <sub>CH</sub>
no. of residues		
Amino acid		
Asp	51.9	50.4
Thr	32.8	30.9
Ser	32.6	33.7
Glu	53.8	56.1
Pro	17.6	20.4
Gly	43.3	43.1
Ala	28.6	30.1
Cys/2	24.0	26.5
Val	32.1	31.5
Met	6.3	3.9
Ile	19.9	17.9
Leu	32.2	32.2
Tyr	18.3	18.1
Phe	23.0	19.8
Lys	29.7	29.1
His	10.0	10.7
Arg	22.1	20.5
Trp	10.0	10.1
Gla*	9.8	10.2
Total residues	498.0	495.2
Mol wt	56,600	56,600
%		
Carbohydrates		
Galactose	1.81	1.93
Mannose	1.72	1.65
Glucose	0.31	0.41
Fucose	0.39	0.51
Hexosamines	7.30	7.00
Sialic Acid	6.40	7.60
Total	17.93	19.10

\*  $\gamma$ -Carboxyglutamic acid.

slightly higher sialic acid content of Factor IX<sub>CH</sub>, but this is within the limits of experimental error.

The results of activation of both normal Factor IX and Factor IX<sub>CH</sub> are shown in Fig. 4. In this experiment, when normal Factor IX is activated by Factor XIa and Ca<sup>2+</sup>, the clotting times are progressively shortened. When purified Factor IX<sub>CH</sub> is activated by Factor XI and Ca<sup>2+</sup> under the same conditions, very little shortening of the unactivated partial thromboplastin time is observed. If Factor XIa, Factor IX, or Ca<sup>2+</sup> is omitted from the incubation mixture as a control, there is again very little shortening of the clotting time.

When the activation of normal Factor IX is followed by SDS gel electrophoresis, as shown in Fig. 5, the 70,000-mol wt band is converted to a lower molecular weight component (estimated mol wt 50,000) within 30 min. When Factor IX<sub>CH</sub> is activated under the same conditions, no change in molecular weight is observed. Even after incubation for 20 h, only very little lower molecular weight component was formed. These findings were consistent on several determinations.

The extinction coefficients ( $E_{280}^{1\%}$ ) were similar, 11.8 for Factor IX and 12.5 for Factor IX<sub>CH</sub>.

## DISCUSSION

An abnormal Factor IX molecule obtained from a patient with mild hemophilia B has been isolated and

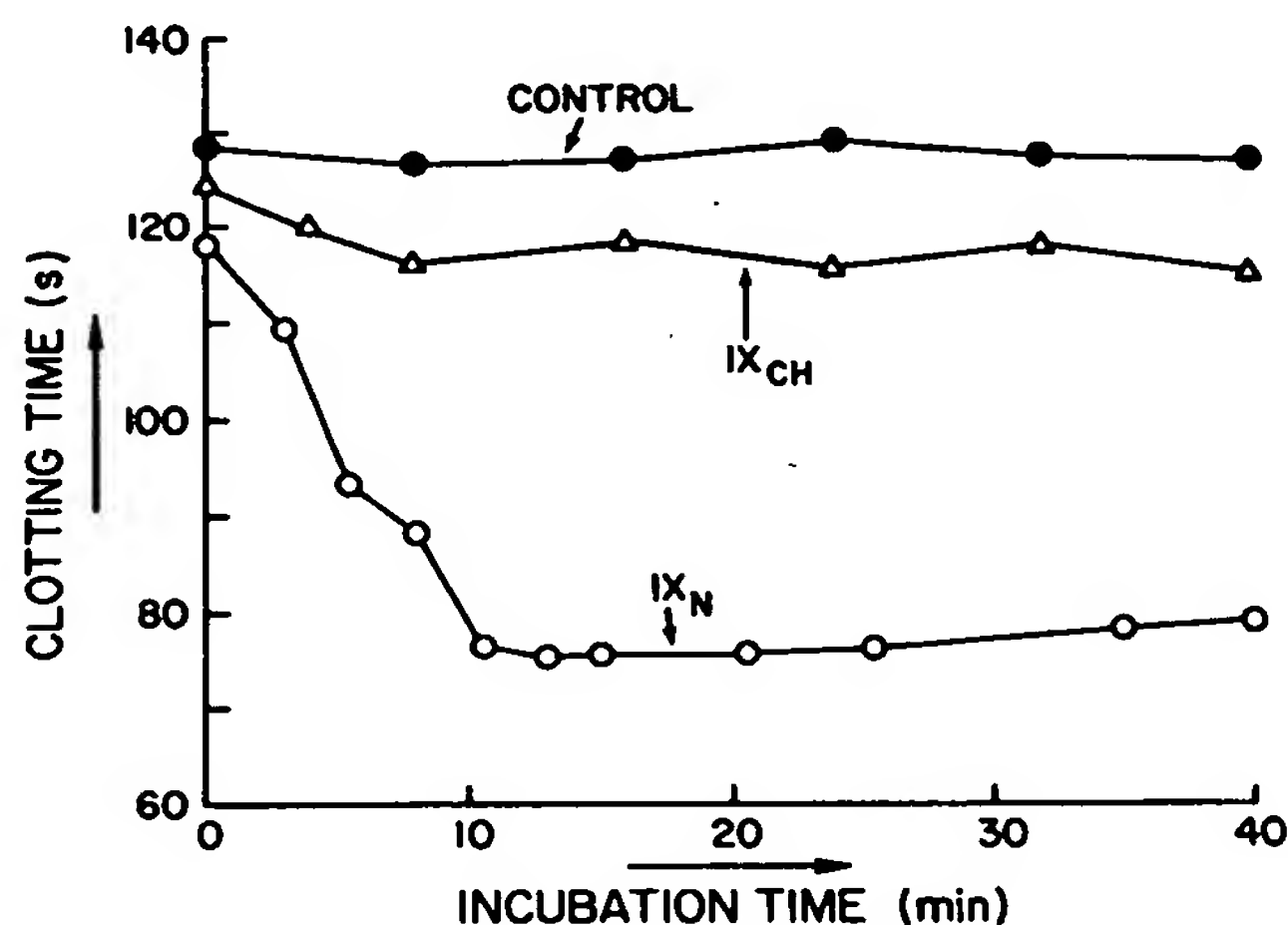


FIGURE 4 Activation of Factor IX<sub>CH</sub> and normal human Factor IX by activated human Factor XI and Ca. The incubation mixture consisted of 3 mg Factor IX protein (or buffer control); 100  $\mu$ g Factor XIa (or buffer control); 5 mM CaCl<sub>2</sub> in a final volume of 3 ml (0.05 M imidazole buffer, pH 6.0). At the indicated time intervals, 0.1 ml of the incubation mixture was diluted with imidazole buffer (1:200 for Factor IX and 1:20 for Factor IX<sub>CH</sub>) and was added to 0.1 ml unactivated Factor IX-deficient substrate. Clotting was achieved by the addition of 0.2 ml of a 1:1 mixture of 32 mM CaCl<sub>2</sub> and 10% Thrombofax in normal saline.

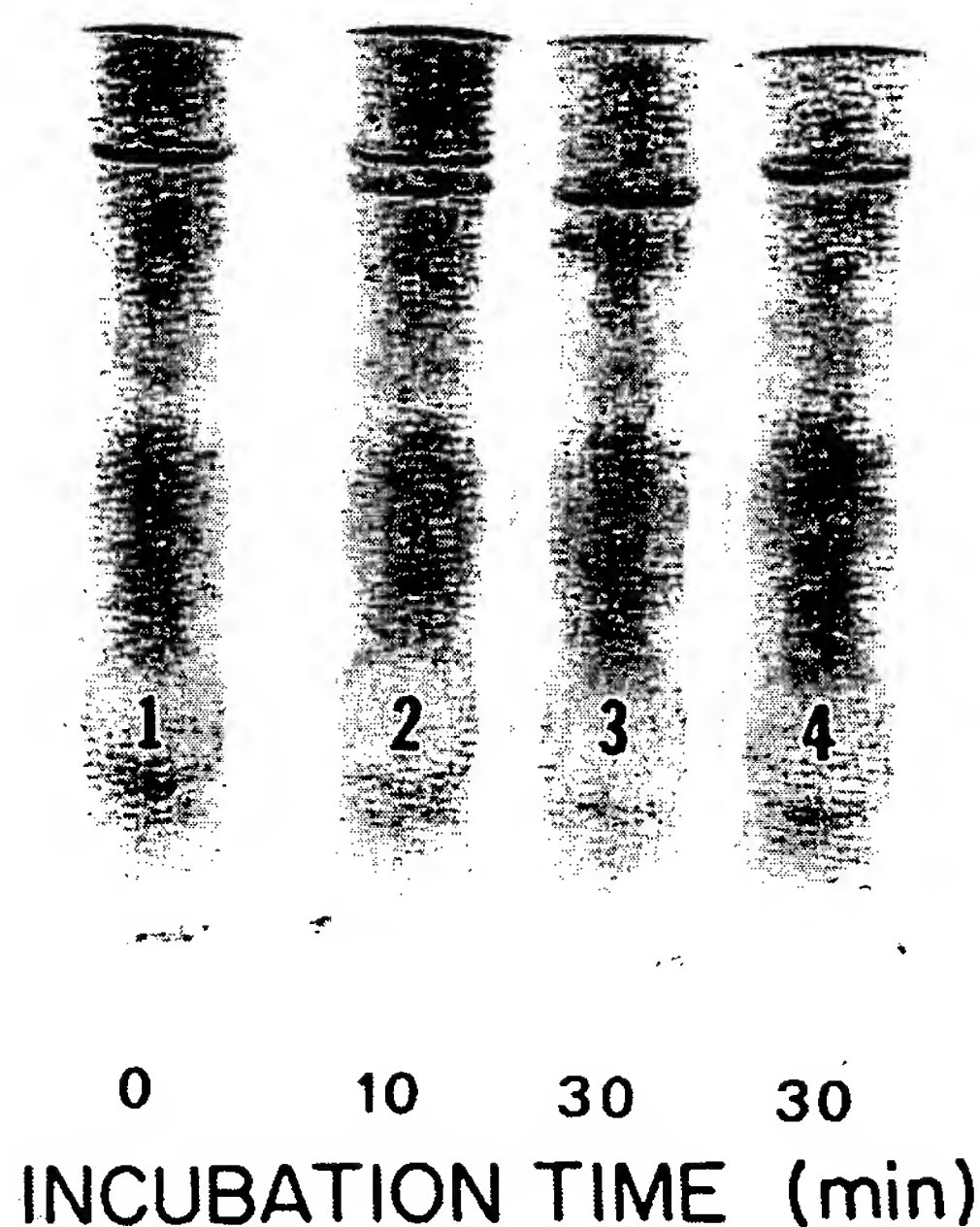


FIGURE 5 Activation of Factor IX<sub>CH</sub> and normal human Factor IX by activated human Factor XI and Ca. The activation conditions were the same as those described in Fig. 4 except that at the indicated time interval, an aliquot of the incubation mixture which contained  $\approx 15 \mu\text{g}$  protein was subjected to SDS gel electrophoresis. Gels 1-3 show normal Factor IX which was converted to a lower molecular weight component. Gel 4 shows Factor IX<sub>CH</sub> which showed no change in molecular weight.

characterized. The patient's plasma contains 5% Factor IX clotting activity but 100% Factor IX antigen as determined by radioimmunoassay and by antibody neutralization experiments as previously described. This variant Factor IX (Factor IX<sub>CH</sub>) was first described by Aronson (23) who noted that addition of SrCl<sub>2</sub> to the patient's plasma did not result in activation of the patient's Factor IX, while addition of SrCl<sub>2</sub> to normal plasma caused rapid activation of normal Factor IX. After activation promoted by SrCl<sub>2</sub>, the apparent molecular weight of the patient's Factor IX did not change, while that of normal Factor IX, as estimated on a Sephadex column (Pharmacia Fine Chemicals, Inc.), had a smaller molecular weight (23). However, Aronson used a plasma preparation, not a purified Factor IX.

We have purified Factor IX<sub>CH</sub> 18,000-fold. The product is homogeneous by several criteria. In the purification procedure, the DE-52 ion exchange and the heparin-agarose affinity chromatography are critical. In the former, the flow rate must be maintained at least at 60 ml/h. On heparin-agarose chromatography, calcium ions were used to enhance the binding of both normal Factor IX and Factor IX<sub>CH</sub> to the column. In

the presence of 2.5 mM CaCl<sub>2</sub>, prothrombin does not bind to the column and can be collected in relatively pure form after the sample application. The column is washed extensively with imidazole buffer, which contains 0.25 M NaCl and 2.5 mM CaCl<sub>2</sub>, to remove Factor X and prothrombin remaining in the column. Factor IX<sub>CH</sub> can then be eluted with 0.4 M NaCl in the same buffer.

Factor IX<sub>CH</sub> resembles normal human Factor IX in several respects. Although several other values of molecular weight for normal human Factor IX have been reported (2-6, 8) including one as low as 57,000 as determined by sedimentation equilibrium (8), higher molecular weights are found with the Laemmli SDS gel method (35). Nevertheless, it is apparent that both Factor IX moleculars have a similar molecular weight. In addition, both normal Factor IX and Factor IX<sub>CH</sub> have tyrosine as the NH<sub>2</sub>-terminal amino acid. Amino acid analyses of both proteins are very similar. Compared to analyses for normal human Factor IX published by other investigators (2, 3, 6, 8), by normalizing molecular weights to 50,000, our analyses were found to be in good agreement with some reports in terms of most of the individual amino acid (3, 8) as well as the total residues (8). This type of analysis, however, obviously does not exclude a single amino acid substitution in Factor IX<sub>CH</sub>. Furthermore, the carbohydrate content of the normal and abnormal molecules is similar and is in agreement with that reported by others for the normal molecule (3, 6, 8). Moreover, Factor IX<sub>CH</sub> contains the same number of  $\gamma$ -carboxyglutamic acid residues as normal Factor IX. Thus, there is no evidence that the abnormality in Factor IX<sub>CH</sub> is the result of defective  $\gamma$ -carboxylation of the molecule during postribosomal modification in the hepatic parenchymal cells.

Of particular interest are the studies on the activation of Factor IX<sub>CH</sub>. Activation of this abnormal Factor IX by Factor XIa and Ca<sup>2+</sup>, when followed by both coagulation assays and SDS gel electrophoreses, strongly suggests that the major abnormality in Factor IX<sub>CH</sub> is defective activation. These findings are in agreement with those reported by Aronson (23), except that activation of the purified protein is even slower than in a plasma system. Thus, it is possible that the major defect of Factor IX<sub>CH</sub> is delayed activation under physiological conditions. Factor IX<sub>CH</sub> may be the result of a single amino acid substitution governed by gene(s) on the X chromosome. Such a substitution could retard the cleavage of bonds that result in the two-chain Factor IXa molecule. Prolonged incubation of Factor IX<sub>CH</sub> with Factor XIa and Ca<sup>2+</sup> for up to 20 h does result in the formation of a lower molecular weight component, but this process is very slow and most of the Factor IX<sub>CH</sub> remains in the zymogen form. Activated Factor IX<sub>CH</sub> has not yet been isolated in amounts sufficient to com-

pare specific activity of Factor IX<sub>CH</sub> with that of normal Factor IXa. Fujikawa et al. (10) and Lindquist et al. (48) have shown, with purified bovine Factor IX, that an Arg-Ala bond is selectively cleaved by Factor XIa resulting in an inactive intermediate; a second proteolytic step cleaves an Arg-Val bond, releases a carbohydrate-rich activation peptide, and results in activated Factor IX. Whether the amino acid substitution in Factor IX<sub>CH</sub> is close to sites of cleavage by Factor XIa and Ca<sup>2+</sup>, or whether the substitution is remote and exerts its effect through tertiary conformational changes that hinder the proteolytic action of Factor XIa is not yet elucidated.

While it seems clear that Factor IX<sub>CH</sub> is not activated normally, it is difficult to determine whether this is the only defect. If Factor IX<sub>CH</sub> can be activated by substances such as Russell's viper venom or kallikrein, it would be possible to study the interaction of Factor IX<sub>CH</sub> with calcium, lipid, and Factors VIII and X. These studies are currently in progress in our laboratory.

It could be postulated that the patient with Factor IX<sub>CH</sub> is synthesizing both a normal protein (that is contributing the 5% coagulant activity) and a nonfunctioning protein that retains antigenic activity but not clotting activity. This explanation seems unlikely on a genetic basis because hemophilia B is inherited as an X-linked recessive characteristic and it would be unlikely to have a gene coding for normal Factor IX and another for abnormal Factor IX. Moreover, in the activation experiment with clotting assays, one would expect 5% normal Factor IX to activate normally which was not the case. The Factor IX protein purified from the affected patient is probably composed entirely of a Factor IX protein that does not function normally in coagulation reactions.

Other abnormal Factor IX molecules almost certainly exist as initially postulated by Fantl et al. (49), Roberts et al. (12), and Hougie and Twomey (50). The latter authors described a hemophilia B<sub>M</sub> kindred which was shown to be CRM<sup>+</sup> by Brown et al. (16) and which has recently been purified and characterized by Østerud et al. (51). Factor IX<sub>BM</sub> differs from Factor IX<sub>CH</sub> in that the former is associated with a prolonged ox-brain prothrombin time and can be activated by Factor XIa in the presence of Ca<sup>2+</sup> (51). Furthermore, it should be noted that Factor IX<sub>CH</sub> is different from the Factor IX abnormality in hemophilia B Leyden (52). The Factor IX level in hemophilia B Leyden increases as the patient gets older. This is not the case with the patient with Factor IX<sub>CH</sub> who has been followed since childhood. Factor IX<sub>CH</sub> is also different from CRM<sup>R</sup> and CRM<sup>-</sup> kindred previously reported by us (53). Furthermore, Factor VII levels are normal in the patient with Factor IX<sub>CH</sub> who is thus different from the Factor IX variants reported by Kasper et al. (14) and Girolami et al. (15).

Studies of abnormal Factor IX molecules should help

in resolving some of the structure function relationships of normal Factor IX because the putative amino acid substitutions in the abnormal forms may affect different functions of the Factor IX molecule. For example, some abnormal Factor IX proteins may exhibit delayed activation by Factor XIa and calcium, while others may exhibit defective interaction with other blood clotting factors. The molecular basis of the interaction of Factor IX with other blood clotting factors, calcium, and phospholipid should provide better insight into the pathophysiology of the clinical disease state, hemophilia B.

## ACKNOWLEDGMENTS

We are indebted to Doctors Richard Hiskey and Roger Lundblad for helpful suggestions and criticisms. We also wish to thank Mrs. K. Su Chung, Miss Patti Bean, and Mrs. Thelma Duncan for expert technical and secretarial assistance.

This work was supported by United States Public Health grants HL 06350, HL 07149, HL 16633, and HL 20161 from the National Heart, Lung, and Blood Institute.

## REFERENCES

1. Fujikawa, K., A. R. Thompson, M. E. Legaz, R. G. Meyer, and E. W. Davie. 1973. Isolation and characterization of bovine factor IX (Christmas factor). *Biochemistry*. 12: 4938-4945.
2. Østerud, B., and R. Fløengsrud. 1975. Purification and some characteristics of the coagulation factor IX from human plasma. *Biochem. J.* 145: 469-474.
3. Andersson, L.-O., H. Borg, and M. Miller-Andersson. 1975. Purification and characterization of human factor IX. *Thromb. Res.* 7: 451-459.
4. Rosenberg, J. S., P. W. McKenna, and R. D. Rosenberg. 1975. Inhibition of human factor IX<sub>a</sub> by human antithrombin. *J. Biol. Chem.* 250: 8883-8888.
5. Thompson, A. R. 1977. Factor IX antigen by radioimmunoassay. Abnormal factor IX protein in patients on warfarin therapy and with hemophilia B. *J. Clin. Invest.* 59: 900-910.
6. Suomela, H. 1976. Human coagulation factor IX: isolation and characterization. *Eur. J. Biochem.* 71: 145-154.
7. Chung, K. S., H. M. Reisner, and H. R. Roberts. 1976. Purification and characterization of human blood clotting factor IX. *Circulation*. 54(II): 118. (Abstr.)
8. Di Scipio, R. G., M. A. Hermanson, S. G. Yates, and E. W. Davie. 1977. A comparison of human prothrombin, factor IX (Christmas factor), factor X (Stuart factor), and protein S. *Biochemistry*. 16: 698-706.
9. Kingdon, H. S., and R. L. Lundblad. 1975. Biochemistry of factor IX. In *Handbook of Hemophilia*. K. M. Brinkhous and H. C. Hemker, editors. Excerpta Medica, Amsterdam, The Netherlands. 103-118.
10. Fujikawa, K., M. E. Legaz, H. Kato, and E. W. Davie. 1974. The mechanism of activation of bovine factor IX (Christmas factor) by bovine factor XI<sub>a</sub> (activated plasma thromboplastin antecedent). *Biochemistry*. 13: 4508-4516.
11. Titani, K., D. L. Enfield, K. Katayama, L. H. Ericsson, K. Fujikawa, K. A. Walsh, and H. Neurath. 1977. Primary structure of bovine factor IX. *Thromb. Haemostasis*. 38: 116. (Abstr.)
12. Roberts, H. R., J. E. Grizzle, W. D. McLester, and G. D. Penick. 1968. Genetic variants of hemophilia B: Detec-



- tion by means of a specific PTC inhibitor. *J. Clin. Invest.* 47: 360-365.
13. Ørstavik, K. H., B. Østerud, H. Prydz, and K. Berg. 1975. Electroimmunoassay of factor IX in hemophilia B. *Thromb. Res.* 7: 373-382.
  14. Kasper, C. K., B. Østerud, J. Y. Minami, W. Shonick, and S. I. Rapaport. 1977. Hemophilia B: Characterization of genetic variants and detection of carriers. *Blood.* 50: 351-366.
  15. Girolami, A., A. Sticchi, A. Burul, and R. Dal Bo Zanon. 1977. An immunological investigation of hemophilia B with a tentative classification of the disease into five variants. *Vox. Sang.* 32: 230-238.
  16. Brown, P. E., C. Hougie, and H. R. Roberts. 1970. The genetic heterogeneity of hemophilia B. *N. Engl. J. Med.* 283: 61-64.
  17. Meyer, D., E. Bidwell, and M. J. Larrieu. 1972. Cross reacting material in genetic variants of haemophilia B. *J. Clin. Pathol. (Lond.)* 25: 433-436.
  18. Pfueller, S., J. B. Somer, and P. A. Castaldi. 1969. Haemophilia B due to an abnormal factor IX. *Coagulation.* 2: 213-219.
  19. Elödi, S., and E. Puskas. 1972. Variants of hemophilia B. *Thromb. Diath. Haemorrh.* 28: 489-495.
  20. McLester, W. D., H. R. Roberts, and R. H. Wagner. 1965. Use of immunosorbent technique in the study of a PTC inhibitor: A new method for the investigation of blood coagulation. 1965. *J. Lab. Clin. Med.* 66: 682-687.
  21. Roberts, H. R., G. P. Gross, W. P. Webster, I. I. Dejanov, and G. D. Penick. 1966. Acquired inhibitors of plasma factor IX: A study of their induction, properties and neutralization. *Am. J. Med. Sci.* 251: 43-50.
  22. Neal, W. R., D. T. Tayloe, Jr., A. I. Cederbaum, and H. R. Roberts. 1973. Detection of genetic variants of hemophilia B with an immunosorbent technique. *Br. J. Haematol.* 25: 63-68.
  23. Aronson, D. L. 1975. Recent advances in hemophilia. *Ann. N. Y. Acad. Sci.* 240: 92-94.
  24. Chung, K. S., J. C. Goldsmith, and H. R. Roberts. 1976. Purification and characterization of factor IX Chapel Hill. *Blood.* 48: 974. (Abstr.)
  25. Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall. 1951. Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* 193: 265-275.
  26. Langdell, R. D., R. H. Wagner, and K. M. Brinkhous. 1953. Effect of antihemophilic factor on one-stage clotting tests: a presumptive test for hemophilia and a simple one-stage anti-hemophilic factor assay procedure. *J. Lab. Clin. Med.* 41: 637-647.
  27. Barrow, E. M., W. Bullock, and J. B. Graham. 1960. A study of the carrier state for plasma thromboplastin component (PTC, Christmas factor) deficiency, utilizing a new assay procedure. *J. Lab. Clin. Med.* 55: 936-945.
  28. Hougie, C., E. M. Barrow, and J. B. Graham. 1957. Stuart clotting defect. I. Segregation of an hereditary hemorrhagic state from the heterogeneous group heretofore called "Stable Factor" (SPCA, proconvertin, Factor VII) deficiency. *J. Clin. Invest.* 36: 485-496.
  29. Smith, H. P., E. D. Warner, and K. M. Brinkhous. 1937. Prothrombin deficiency and the bleeding tendency in liver injury (Chloroform intoxication). *J. Exp. Med.* 66: 801-811.
  30. Horowitz, H. I., W. P. Wilcox, and M. M. Fujimoto. 1963. Assay of plasma thromboplastin antecedent (PTA) with artificially depleted normal plasma. *Blood.* 22: 35-43.
  31. Kingdon, H. S., E. W. Davie, and O. D. Ratnoff. 1964. The reaction between activated plasma thromboplastin antecedent and diisopropylphosphorofluoridate. *Biochemistry.* 3: 166-173.
  32. Reisner, H. M., K. S. Chung, and H. R. Roberts. 1976. A radioimmunoassay for factor IX utilizing human inhibitors. American Society of Hematology Annual Session. Grune & Stratton, Inc., New York. 397. (Abstr.)
  33. Porath, J. 1974. General methods and coupling procedures. *Methods Enzymol.* 34: 13-30.
  34. Cuatrecasas, P. 1970. Protein purification by affinity chromatography. *J. Biol. Chem.* 245: 3059-3065.
  35. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (Lond.)* 227: 680-685.
  36. Lundblad, R. L. and H. S. Kingdon. 1974. Biochemistry of the interaction of bovine factors XIa and IX. *Thromb. Diath. Haemorrh.* 57: 315-333.
  37. Edman, P., and G. Begg. 1967. A protein sequenator. *Eur. J. Biochem.* 1: 80-91.
  38. Woods, K. R., and K. T. Wang. 1967. Separation of dansyl-amino acids by polyamide layer chromatography. *Biochim. Biophys. Acta.* 133: 369-370.
  39. Spackman, D. H., W. H. Stein, and S. Moore. 1958. Automatic recording apparatus for use in the chromatography of amino acids. *Anal. Chem.* 30: 1190-1206.
  40. Hauschka, P. V., J. B. Lian, and P. M. Gallop. 1975. Direct identification of the calcium-binding amino acid,  $\gamma$ -carboxyglutamate in mineralized tissue. *Proc. Natl. Acad. Sci. U. S. A.* 72: 3925-3929.
  41. Moore, S. 1963. On the determination of cystine as cysteic acid. *J. Biol. Chem.* 238: 235-237.
  42. Benize, W. L., and K. Schmid. 1957. Determination of tyrosine and tryptophan in proteins. *Anal. Chem.* 29: 1193-1196.
  43. Niedermeier, W. 1971. Gas chromatography of neutral and amino sugars in glycoproteins. *Anal. Chem.* 40: 465-475.
  44. Laine, R., H. Söderlund, and O. Renkonen. 1973. Chemical composition of Semliki Forest virus. *Intervirology.* 1: 110-118.
  45. Dubois, M., K. A. Gilles, J. K. Hamilton, P. A. Rebers, and F. Smith. 1956. Colorimetric method for determination of sugars and related substances. *Anal. Chem.* 28: 350-356.
  46. Gardell, S. 1957. Determination of hexosamines. *Methods Biochem. Anal.* 6: 289-317.
  47. Warren, L. 1959. The thiobarbituric acid assay of sialic acid. *J. Biol. Chem.* 234: 1971-1975.
  48. Lindquist, P. A., K. Fujikawa, and E. W. Davie. 1978. Activation of bovine factor IX (Christmas factor) by factor XIa (activated plasma thromboplastin antecedent) and a protease from Russell's viper venom. *J. Biol. Chem.* 253: 1902-1909.
  49. Fantl, P., R. J. Sawers, and A. G. Marr. 1956. Investigation of a haemorrhagic disease due to beta-prothromboplastin deficiency complicated by a specific inhibitor of thromboplastin formation. *Australas. Ann. Med.* 5: 163-176.
  50. Hougie, C., and J. J. Twomey. 1967. Haemophilia B<sub>M</sub>: A new type of factor IX deficiency. *Lancet.* i: 698-700.
  51. Østerud, B., K. Lavine, C. K. Kasper, and S. I. Rapaport. 1977. Isolation and properties of the abnormal factor IX molecule of hemophilia B<sub>M</sub>. *Thromb. Haemostasis.* 38: 51. (Abstr.)
  52. Veltkamp, J. J., J. Meilof, H. G. Remmelts, D. van der Vlerk, and E. A. Loeliger. 1970. Another genetic variant of haemophilia B: Haemophilia B Leyden. *Scand. J. Haematol.* 7: 82-90.
  53. Roberts, H. R., and A. I. Cederbaum. 1975. Molecular variants of factor IX. In *Handbook of Hemophilia*. K. M. Brinkhous and H. C. Hemker, editors. Excerpta Medica, Amsterdam, The Netherlands. 237-246.

## Multiple Extracellular Domains of CCR-5 Contribute to Human Immunodeficiency Virus Type 1 Entry and Fusion

LAURENT PICARD,<sup>1\*</sup> GRAHAM SIMMONS,<sup>1</sup> CHRISTINE A. POWER,<sup>2</sup> ALEXANDRA MEYER,<sup>2</sup>  
ROBIN A. WEISS,<sup>1</sup> AND PAUL R. CLAPHAM<sup>1</sup>

*Chester Beatty Laboratories, Institute of Cancer Research, London SW3 6JB, United Kingdom,<sup>1</sup> and Geneva Biomedical Research Institute, GlaxoWellcome Research and Development S.A., 1228 Plan-les-Ouates, Geneva, Switzerland<sup>2</sup>*

Received 25 February 1997/Accepted 14 April 1997

Human immunodeficiency virus type 1 (HIV-1) entry is governed by the interaction of the viral envelope glycoprotein (Env) with its receptor. The HIV-1 receptor is composed of two molecules, the CD4 binding receptor and a coreceptor. The seven-membrane-spanning chemokine receptor CCR-5 is one of the coreceptors used by primary isolates of HIV-1. We demonstrate that the mouse homolog of CCR-5 (mCCR-5) does not function as an HIV-1 coreceptor. A set of chimeras of human CCR-5 and mCCR-5 was studied for Env-induced cell fusion and HIV-1 infection. Using the HIV-1<sub>ADA</sub> envelope glycoprotein in a syncytium formation assay, we show that replacement of any fragment containing extracellular domains of mCCR-5 by its human counterparts is sufficient to allow Env-induced fusion. Conversely, replacement of any fragment containing human extracellular domains by its murine counterpart did not lead to coreceptor function loss. These results show that several domains of CCR-5 participate in coreceptor function. In addition, using a panel of primary nonsyncytium-inducing and syncytium-inducing isolates that use CCR-5 or both CXCR-4 and CCR-5 as coreceptors, we show that the latter dual-tropic isolates are less tolerant to changes in CCR-5 than strains with a more restricted coreceptor use. Thus, different strains are likely to have different ways of interacting with the CCR-5 coreceptor.

Human immunodeficiency virus type 1 (HIV-1) enters target cells by fusion of its envelope with the membrane of the cell. This process is initiated by the high-affinity binding of the envelope glycoproteins (Env and gp120) to the CD4 molecule. Subsequent to the Env-CD4 interaction, a number of conformational changes in gp120 and probably in CD4 take place, leading to virus entry or syncytium formation (reviewed in references 58 and 83). These post-CD4 binding events are not completed in most nonhuman cell lines and in some human cell lines expressing human CD4, and membrane fusion fails to occur (5, 20, 23, 53). It was suggested that these cells lack a component necessary for the fusion and entry processes to be completed (13, 35, 41). Many of these experiments were done with syncytium-inducing (SI) T-cell-line-adapted (TCLA) strains of HIV-1, but later evidence suggested that different molecules may be required for the entry of primary non-SI (NSI) isolates (2, 12). The second components of the receptor complex, or coreceptors, have been identified (3, 22, 32, 34, 36, 38) and belong to the seven-transmembrane G protein-coupled chemokine receptor family (reviewed in reference 69).

As predicted, the coreceptors for TCLA isolates and primary isolates are different: the CXCR-4 chemokine receptor (previously known as LCR1, LESTR, HUMSTR, or fusin) functions as an HIV-1 coreceptor for TCLA as well as for primary SI or T-cell-tropic strains, whereas all primary NSI or macrophage-tropic (M-tropic) strains tested so far use the CCR-5 chemokine receptor (3, 22, 32, 34, 36, 38, 78). Other members of the chemokine receptor family, including CCR-2b and CCR-3, can also be used by a restricted number of HIV-1

isolates (22, 34, 42, 71). This extension of the use of the coreceptor by primary isolates has been shown to correlate with disease progression (30). However, CXCR-4 and CCR-5 seem to be the main coreceptors used by HIV-1 in vivo, irrespective of the viral genetic subtype (89). The importance of CCR-5 as an HIV-1 coreceptor in vivo has been confirmed with the identification of a CCR-5-defective allele in some multiply exposed uninfected individuals (31, 43, 51, 73).

Chemokines are small proteins involved in cellular recruitment and activation. They show some sequence homology and can be divided into four subgroups according to the spacing of their amino-terminal cysteine residues. The two main groups are CXC (or  $\alpha$ ) chemokines and CC (or  $\beta$ ) chemokines (reviewed in references 69 and 84). CXCR-4 binds the CXC chemokine stromal cell-derived factor 1 (9, 60), and CCR-5 binds RANTES, macrophage inflammatory protein 1 $\alpha$  (MIP-1 $\alpha$ ), and MIP-1 $\beta$  (29, 70, 72), which are CC chemokines. These chemokine ligands are able to block HIV-1 entry and fusion (9, 26, 27, 60, 63, 65). In addition, receptor antagonists based on these chemokines can also block HIV-1 entry (4, 76). The exact mechanism of this inhibition is not yet known but most likely involves steric blockade of the coreceptor by the ligand. Alternatively, chemokine-mediated inhibition of HIV entry may be due to down-regulation of the receptor on the surfaces of permissive cells.

In fact, it has recently been suggested that a ternary complex consisting of gp120, CD4, and the coreceptor forms on the surfaces of permissive cells (50, 81, 88). The molecular details of the interaction between gp120, CD4, and the coreceptor are still not clear. It has been demonstrated that the third hyper-variable region (V3 loop) of the gp120 envelope glycoprotein is likely to be a component of the coreceptor binding site (81, 88) and to influence coreceptor choice (22, 33, 67), but the interaction between gp120 and the coreceptor molecule has not yet

\* Corresponding author. Mailing address: Chester Beatty Laboratories, Institute of Cancer Research, 237 Fulham Rd., London SW3 6JB, United Kingdom. Phone: 44 171 352 8133. Fax: 44 171 352 3299. E-mail: laurent@icr.ac.uk.



been completely characterized. Using molecular chimeras composed of two members of the CC chemokine receptor family, CCR-5 and CCR-2b (6, 71), or of mouse and human CCR-5 (mCCR-5 and hCCR-5, respectively) (6, 8), three groups of researchers have recently shown that the amino-terminal and possibly the first and second extracellular loops of CCR-5 are important for coreceptor function and specificity. Two of these studies, however, assessed coreceptor use by cell-cell fusion only (8, 71), while the third (6) used cell-free virus infection, but with a limited number of isolates only. Although cell-cell fusion usually correlates well with cell-free virus infection, some exceptions to this rule have been reported (55, 77).

To map the regions important for CCR-5 to function as an HIV-1 coreceptor for cell-cell fusion and virion-cell fusion, we took advantage of the existence of the mouse homolog of CCR-5 (mCCR-5) (10, 57). We confirm here that this molecule does not function as a coreceptor for HIV-1. By generating molecular chimeras composed of hCCR-5 and mCCR-5, we have further defined regions in the coreceptor that are important for HIV-1 entry and fusion with Env-expressing cells. Several domains of CCR-5 are shown to be involved in HIV-1 entry. Moreover, different HIV-1 isolates showed various patterns of interaction, implying a differential use of the CCR-5 coreceptor for NSI M-tropic versus SI dual-tropic strains.

#### MATERIALS AND METHODS

**Cells and viruses.** CCC-CD4 cells are cat kidney cells stably transfected with a human CD4 expression vector (24). HeLa-CD4-LTRlacZ (HeLa-P4) cells are HeLa cervical epithelial cells stably transfected with human CD4 and a long terminal repeat (LTR)-lacZ construct (25). Upon infection by HIV or fusion with Tat-expressing cells, transcription of the lacZ gene directed by the LTR is stimulated and leads to the accumulation of  $\beta$ -galactosidase in the cell. HeLa-Env cells are HeLa cells stably transfected with a  $\Delta$ gag-pol provirus containing a *dhfr* gene in place of *nef*, which confers resistance to methotrexate. HeLa-Env<sub>LAI</sub> (74) and HeLa-Env<sub>ADA</sub> (a gift from A. Brelot, Institut Cochin de Génétique Moléculaire, Paris, France) (68) respectively express the HIV-1 LAI and ADA envelope glycoproteins. All these HeLa-Env cells also constitutively express the HIV-1 transactivator Tat. All the adherent cell lines were maintained in Dulbecco modified Eagle medium supplemented with 10% fetal calf serum (FCS) (or 5% FCS for CCC-CD4), antibiotics, and 0.5 mg of G418 (CCC-CD4 and HeLa-P4) per ml or 5  $\mu$ M methotrexate (HeLa-Env).

All the HIV-1 isolates used in this study have been described previously. LAI (82) is a TCLA strain, and GUN-1 (80) is also TCLA but dual tropic, as it can also infect macrophages (56, 77). 89.6, a gift from R. Collman (University of Pennsylvania, Philadelphia), is a primary HIV-1 strain that infects macrophages and certain CD4<sup>+</sup> T-cell lines and is therefore also dual tropic (28). ADA (86), SF-162 (19), M23, and E80 are primary NSI M-tropic strains of HIV-1 (78). All isolates are of subtype B except E80 and M23, which are unclassified. Viral stocks from primary isolates and TCLA strains were produced in peripheral blood mononuclear cells (PBMCs) and H9 cells, respectively, as described previously (78). ADA viral stocks were produced by transfection of the molecular clone in HeLa cells (1). The ADA-producing plasmid, a gift from M. Alizon (Institut Cochin de Génétique Moléculaire, Paris, France), is a recombinant between the 90.1 molecular clone of HIV-1<sub>LAI</sub> and the *KpnI*-*AvaI* fragment encompassing the ADA envelope protein (68).

**hCCR-5 and mCCR-5 chimera construction.** The hCCR-5 (70) (a gift from P. Gray, ICOS Corp., Bothell, Wash.), mCCR-5 (57), and CXCR-4 cDNAs were subcloned by PCR into the pcDNA3 expression vector (Invitrogen, Leek, The Netherlands) between the *HindIII* and *NotI* sites. The sequences, determined by automated sequencing (Applied Biosystems) with a Perkin-Elmer sequencing kit and analyzed with the Genetics Computer Group package, were found to differ at a few positions from the published sequences (Fig. 1). The differences resulted in conservative amino acid (aa) changes: F80L in mCCR-5 and F166L in hCCR-5. The cDNAs encoding the receptors were fused at the carboxy-terminus to the 12-aa-long *c-myc* epitope tag: NH<sub>2</sub>-GGREQLISEEDLA-COOH (37). We decided to tag all the constructs at the end of the C-terminal intracellular domain to avoid interference with potential interactions between CD4 and/or gp120 and extracellular segments of the coreceptor. To generate hCCR-5 and mCCR-5 chimeras, a unique *EcoRV* restriction site was engineered by site-directed mutagenesis (Transformer kit; Clontech, Palo Alto, Calif.) in the hCCR-5 sequence at the same position as in the mCCR-5 sequence (Fig. 1). Using the same technique, we also introduced a unique *KpnI* restriction site in the mCCR-5 sequence at the same position as in the hCCR-5 sequence (Fig. 1). Introduction of these new common sites did not introduce changes in the amino

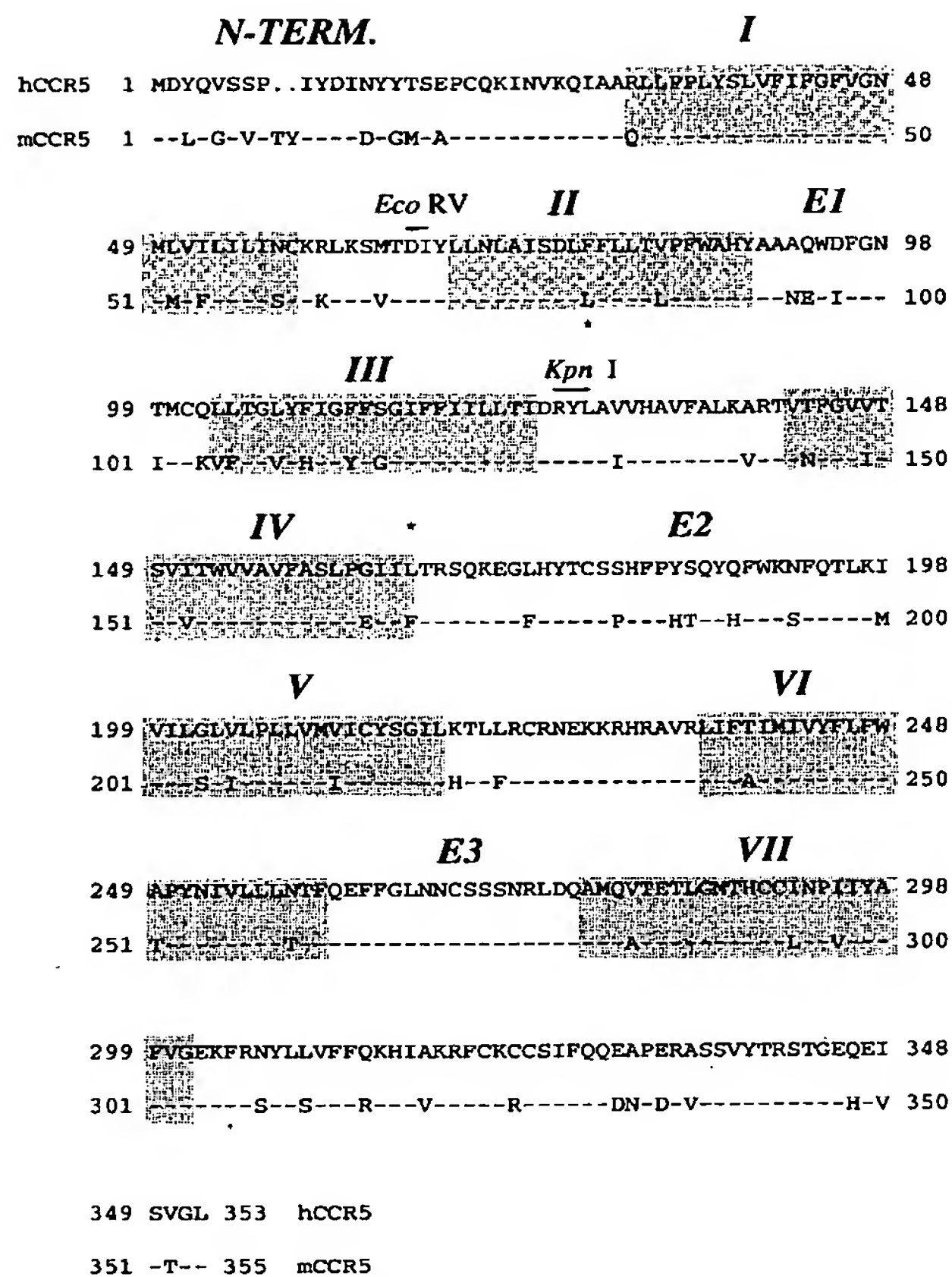


FIG. 1. Amino acid sequence alignment of hCCR-5 and mCCR-5. The two sequences were aligned with the gap program. The sequences display 91% similarity and 83% identity. Potential transmembrane segments (listed under SwissProt accession no. P51682) are shaded and numbered with Roman numerals. Extracellular loops are also numbered. Restriction endonuclease sites are labeled and overlined. Only changes at the amino acid level are shown, while identical amino acids are indicated by dashes. Asterisks identify differences from the published sequences. N-TERM., N terminus.

acid sequences of the proteins. The different chimeric constructs were then generated by exchanging fragments by using the four common unique restriction sites, *HindIII*, *EcoRV*, *KpnI*, and *NotI*. The *HindIII*-*EcoRV* fragment spans the N terminus, transmembrane 1 (TM1), and intracellular loop 1 (I1) of CCR-5; the *EcoRV*-*KpnI* fragment spans TM2, extracellular loop 1 (E1), and TM3; and the *KpnI*-*NotI* fragment spans I2, TM4, E2, TM5, I3, TM6, E3, TM7, and the C terminus. The structures of the hCCR-5-mCCR-5 chimeras are schematically depicted in Fig. 2.

**CCR-5 cell surface expression.** The constructs encoding the different CCR-5 chimeras were transfected into subconfluent HeLa-P4 or CCC-CD4 cells (5  $\mu$ g per 50-mm-diameter petri dish) by the calcium phosphate precipitation method. After 16 to 20 h of incubation, the cells were washed in phosphate-buffered saline (PBS), fed, and incubated for another 48 h. The cells were then fixed in methanol-acetone (1:1) for 5 min at room temperature and subsequently washed in PBS-1% FCS. The cells were then incubated with purified anti-myc (9E10) monoclonal antibody (1:100 in PBS-1% FCS) for 1 h at room temperature. After four washes in PBS-1% FCS, fluorescein isothiocyanate-conjugated goat anti-mouse immunoglobulin G1 antibody (Southern Technologies Associates Inc., Birmingham, Ala.) was added for 1 h at room temperature (1:100 in PBS-1% FCS). After two washes in PBS-1% FCS and two washes in PBS, the cells were mounted in Mowiol (Calbiochem) and observed on a Zeiss fluorescence microscope. Surface expression was estimated by randomly counting 10 fields. Cells were scored as positive if the signal was associated with the cell plasma membrane.



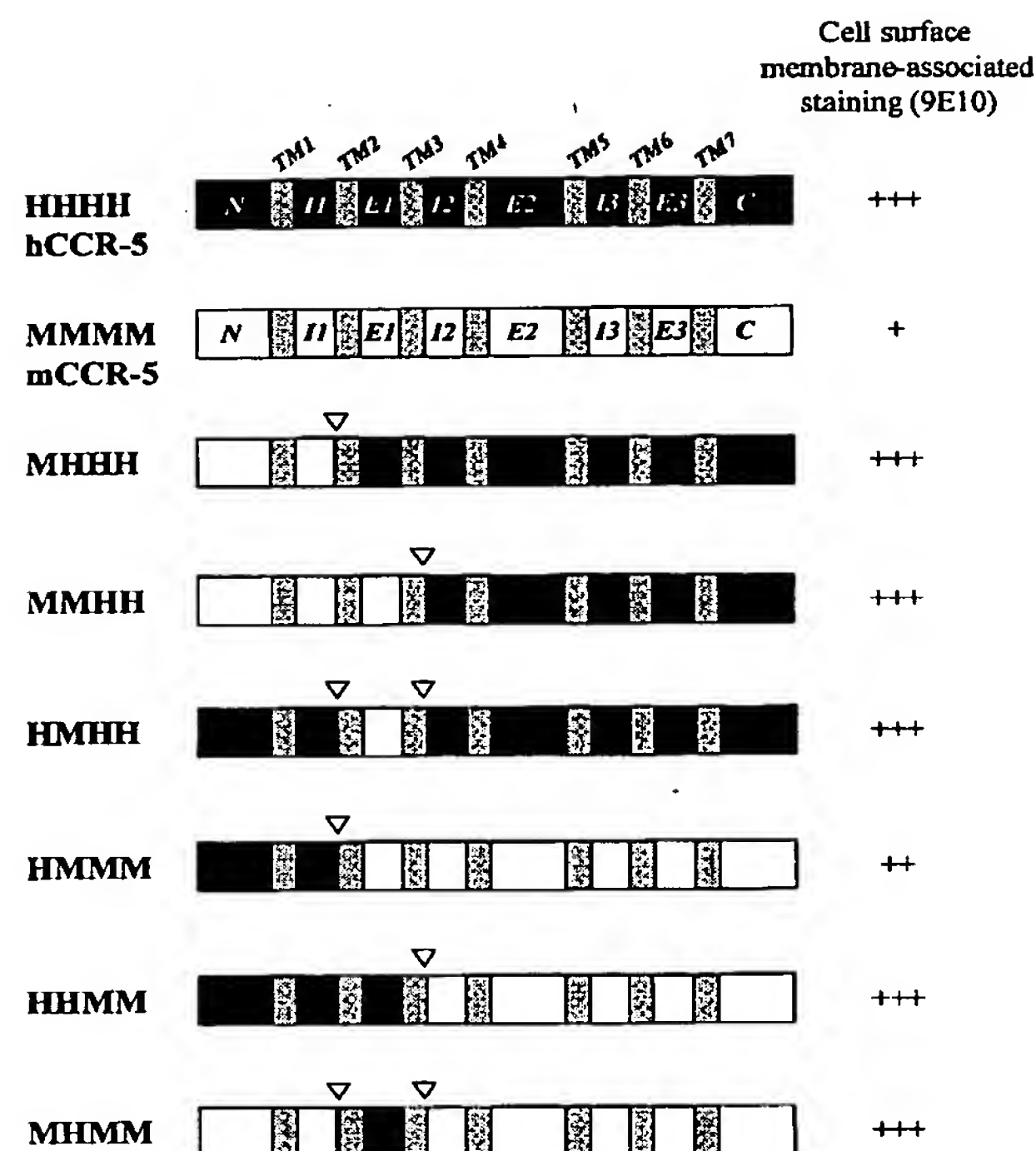


FIG. 2. Schematic representation of hCCR-5-mCCR-5 chimeras. Segments of human origin are filled, while segments of murine origin are open. Transmembrane domains are shaded, and inverted triangles indicate junction sites. The relative level of surface membrane-associated expression of each chimeric receptor, as estimated by immunofluorescence 2 to 3 days after transfection of HeLa-P4 or CCC-CD4 cells (see Materials and Methods), is indicated on the right-hand side and expressed as the percentage of hCCR-5 expression as follows: +, 20 to 50%; ++, 50 to 70%; and +++, 70 to 100%. These estimations are representative of three independent experiments.

**Fusion assay.** HeLa-P4 cells were plated in 12-well (or 24-well) trays and transfected by the calcium phosphate precipitation method the next day with the different CCR-5 constructs with 1 (or 0.5)  $\mu$ g of DNA per well. After 16 to 20 h, the cells were washed in PBS, split into four replicates in 96-well plates, and cocultivated with HeLa-Env<sub>ADA</sub> or HeLa-Env<sub>LAI</sub> ( $10^4$  cells/well) for 24 h. An in situ X-Gal (5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside) assay was then done in duplicate. In brief, cells were washed in PBS, fixed in PBS-0.5% glutaraldehyde for 5 min, washed in PBS, and incubated with the X-Gal substrate (0.5 mg of X-Gal per ml in PBS containing 3 mM ferricyanide, 3 mM ferrocyanide, and 2 mM  $MgCl_2$ ; Calbiochem) for 2 h at 37°C. The two other wells were assayed for  $\beta$ -galactosidase production by a colorimetric assay (59). After being washed in PBS, the cells were lysed in 100  $\mu$ l of PBS-0.5% Nonidet P-40 for 5 min at room temperature. Fifty microliters of that lysate was then mixed with 50  $\mu$ l of 2 $\times$  CPRG substrate (16 mM chlorophenol red- $\beta$ -D-galactopyranoside [Boehringer Mannheim], 0.12 M  $Na_2HPO_4 \cdot 7H_2O$ , 0.08 M  $NaH_2PO_4 \cdot H_2O$ , 0.02 M KCl, 0.02 M  $MgSO_4 \cdot 7H_2O$ , 0.01 M  $\beta$ -mercaptoethanol) in a 96-well plate and incubated for 2 to 3 h in the dark at room temperature. These conditions were checked for linearity of the enzymatic colorimetric reaction. The absorbance at 540 nm was then read on a Dynatech microplate reader.

**Infectivity assay.** Subconfluent CCC-CD4 cells in 100 or 50-mm-diameter petri dishes were transfected with 10 or 5  $\mu$ g of the different CCR-5 constructs by using Lipofectamine (Gibco BRL, Life Technologies) or calcium phosphate precipitation for 5 or 16 to 20 h, respectively. The cells were then split into 48-well plates at  $2 \times 10^4$  to  $4 \times 10^4$  cells/well. They were challenged the next day with HIV-1 ( $\sim 10^3$  to  $10^4$  focus forming units per well) for 2 to 5 h, refed, immunostained 3 days postinfection with an anti-p24 antibody as the primary antibody, and subsequently incubated with a secondary antibody conjugated to  $\beta$ -galactosidase (Genosys Biotechnologies, Inc.) as described previously (24). The blue-stained syncytia were then scored after an in situ X-Gal assay.

## RESULTS

The murine homolog of CCR-5 is inactive as a coreceptor for HIV-1. The murine homolog of CCR-5 (mCCR-5) was recently cloned and has 83% identity and 91% similarity at the amino acid level to hCCR-5 (57). The differences are not located in a particular region of the molecule but are distributed throughout the sequence (Fig. 1). The murine protein is a structural and functional homolog of hCCR-5, since it acts as a receptor for mouse MIP-1 $\alpha$ , mouse MIP-1 $\beta$ , and mouse RANTES (10, 57). Because of its closely related sequence and chemokine binding profile, we decided to test its functionality as an HIV-1 coreceptor.

hCCR-5 and mCCR-5 were tagged at their C termini with the *c-myc* epitope (9E10), cloned into a pcDNA3 expression vector, and transfected into HeLa-P4 cells. This CD4<sup>+</sup> HeLa cell line is stably transfected with a *lacZ* gene under the control of the HIV-1 LTR and is therefore inducible by the HIV-1 transactivator Tat. Because of their human origin and the endogenous expression of the CXCR-4 receptor (38), HeLa-P4 cells are permissive for entry and fusion of TCLA and primary SI HIV-1 strains but not for primary NSI M-tropic strains. After transfection of hCCR-5 or mCCR-5 into HeLa-P4 cells, the cells were cocultivated with HeLa-Env<sub>LAI</sub> or HeLa-Env<sub>ADA</sub> cells. As expected, HeLa-Env<sub>LAI</sub> were able to form syncytia with untransfected cells as well as with hCCR-5- and mCCR-5-transfected cells (not shown). HeLa-Env<sub>ADA</sub>, however, could form syncytia only with HeLa-P4 cells transfected with hCCR-5, as mCCR-5 did not serve as a coreceptor for HIV-1<sub>ADA</sub> Env-induced cell fusion (Fig. 3 and 4). The surface expression level of the mCCR-5 protein in HeLa-P4 cells was fivefold less that of hCCR-5 (Fig. 2). However, overexpression of mCCR-5 in CD4<sup>+</sup> CCC cat kidney cells or 3T3 mouse fibroblasts by vaccinia virus/T7 RNA polymerase infection did not result in syncytium formation upon cocultivation with HeLa-Env<sub>ADA</sub> while hCCR-5 did (data not shown). Thus, despite its similarity with hCCR-5, mCCR-5 is not functional as an HIV-1 coreceptor. We therefore used this property to map the regions of CCR-5 important for HIV-1 entry and fusion by using chimeric CCR-5 receptors.

**Mapping of CCR-5 extracellular domains involved in HIV-1<sub>ADA</sub> envelope-induced fusion.** To map the extracellular domains involved in coreceptor function for HIV-1 NSI M-tropic strains, we chose to use the ADA isolate as an Env prototype. Since hCCR-5 and mCCR-5 are conserved at the amino acid level, it was possible to construct functional molecular chimeras between the two receptors (described in Materials and Methods). The nomenclature that we used for the different regions of chemokine receptors is as follows: E1, E2, and E3 for extracellular loops 1 to 3, amino-terminal extracellular domain, I1, I2, and I3 for intracellular loops 1 to 3, carboxy-terminal intracellular domain, and TM1 to TM7 for the seven-membrane-spanning segments. We reasoned that to be used as an HIV-1 coreceptor, the important functional domains of the molecule were likely to be the four extracellular domains (N terminus, E1, E2, and E3). However, we should bear in mind that the transmembrane and intracellular domains can play a role in the overall conformation of the receptor, in the interaction with gp120 and/or CD4, or in the fusion step per se. There are seven amino acid differences (and a 2-aa insertion in mCCR-5) between hCCR-5 and mCCR-5 amino-terminal extracellular domains (30 to 32 aa long), five differences in E1 (13 aa long), seven differences in E2 (32 aa long), and no differences in E3 (17 aa long). We generated six chimeras by exchanging fragments encompassing each of the first three extracellular domains. The resulting constructs are depicted in

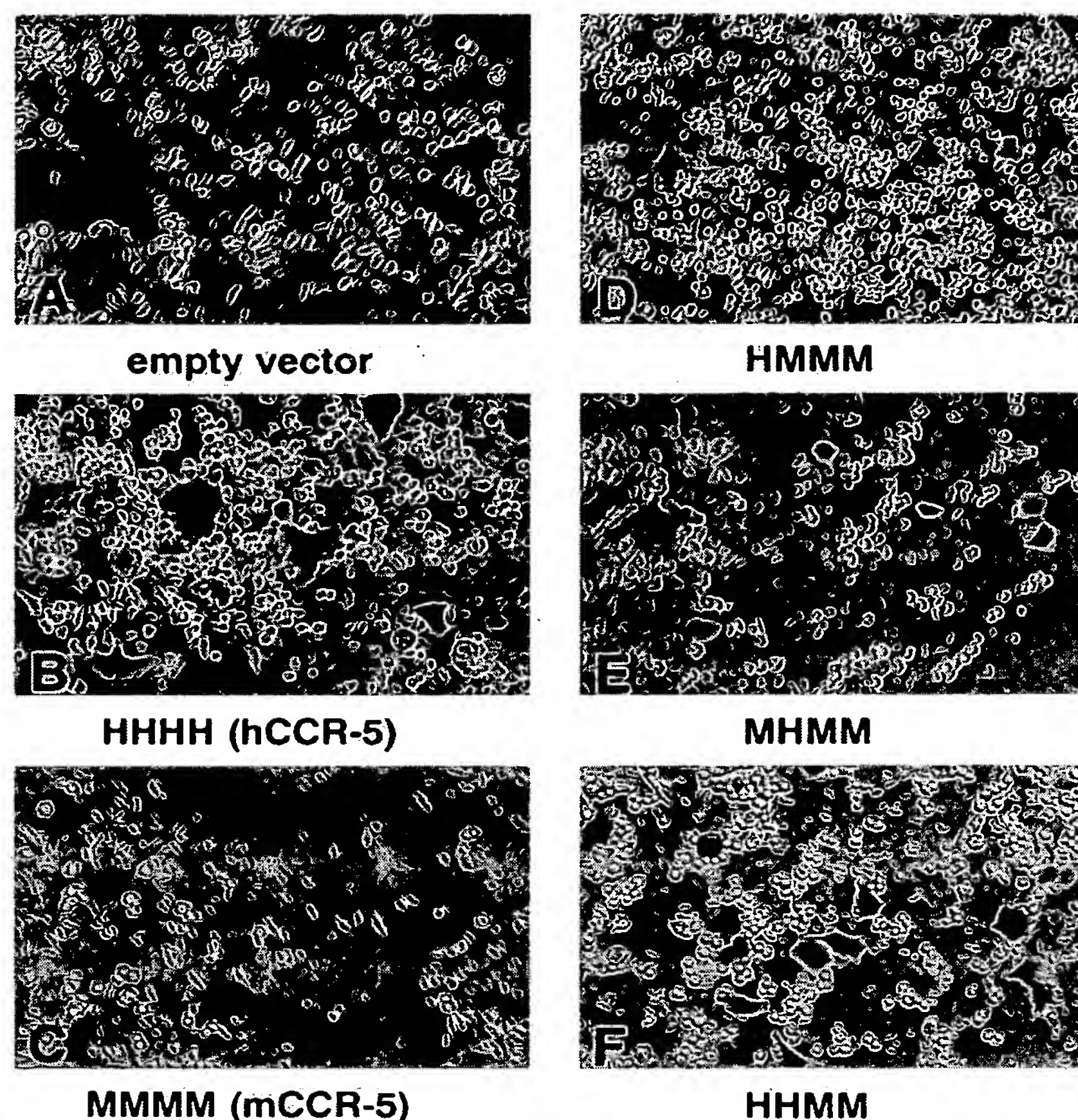


FIG. 3. Formation of syncytia between HIV-1<sub>ADA</sub> envelope-expressing cells and HeLa-P4 cells expressing different hCCR-5-mCCR-5 chimeras. HeLa-P4 cells transiently transfected with an empty vector (A), hCCR-5 (B), mCCR-5 (C), or the different chimeras HMMM (D), MHMM (E), and HHMM (F) were cocultivated for 24 h with HeLa-Env<sub>ADA</sub> cells, fixed, and stained with the X-Gal substrate as described in Materials and Methods. Photographs were taken at a  $\times 160$  magnification. Results to similar those seen in panels B and F were obtained for the MHMH, HMHH, and MMHH chimeric receptors.

Fig. 2. Each of the four extracellular domains is represented by the letter H or M depending on the human or murine origin of the extracellular domains in the chimera, and the constructs are named accordingly.

Levels of cell surface expression of the six chimeras and the two parental clones were estimated by transfecting the C-terminus-tagged constructs into HeLa-P4 or CCC-CD4 cells and staining them with the 9E10 (*c-myc* epitope) monoclonal antibody by indirect immunofluorescence after permeabilization of the cells. Cells were scored as positive if the 9E10 staining was associated with the cell plasma membrane. As previously mentioned, mCCR-5 was expressed at a lower level than its human counterpart. All other hCCR-5-mCCR-5 constructs were expressed at levels similar to that of the hCCR-5 parental clone except for the HMMM chimera, which showed about 50% of the level of hCCR-5 expression (Fig. 2).

These different chimeras were tested for coreceptor activity by transient transfection into HeLa-P4 cells and coculture with HeLa-Env<sub>ADA</sub>. Blue-stained syncytia were scored after an *in situ* X-Gal assay, or  $\beta$ -galactosidase activity was measured by a CPRG colorimetric assay (Fig. 3 and 4). To map the domain(s) necessary and sufficient to confer coreceptor activity to

mCCR-5, we first introduced segments of hCCR-5 containing each extracellular domain into nonfunctional mCCR-5. The HMMM chimera inefficiently mediated HIV-1<sub>ADA</sub> Env-induced fusion (Fig. 3 and 4), indicating that a fragment encompassing the N-terminal, TM1, and I1 domains was not sufficient to confer complete coreceptor activity or that the lower level of expression of this construct resulted in reduced coreceptor activity. However, when we introduced a fragment containing the TM1, E1, and TM2 domains in mCCR-5 (MHMM), it was possible to observe syncytium formation with Env-expressing cells. This fusion activity was reproductively about 30% of the level of the hCCR-5 control when we assayed by the CPRG colorimetric assay (Fig. 4). This reduced activity was due to a reduction in the size of syncytia (Fig. 3). Thus, the first extracellular loop when it is introduced into the mCCR-5 receptor is sufficient to mediate ADA Env-mediated fusion. To determine the role of the second extracellular loop, we used the MMHH chimera, which is identical in extracellular sequence to the theoretical MMHM chimera. This chimera had coreceptor activity similar to that of parental hCCR-5, indicating that E2 is also important in mediating HIV-1 Env interaction.

We then placed two domains of hCCR-5 into the nonfunc-



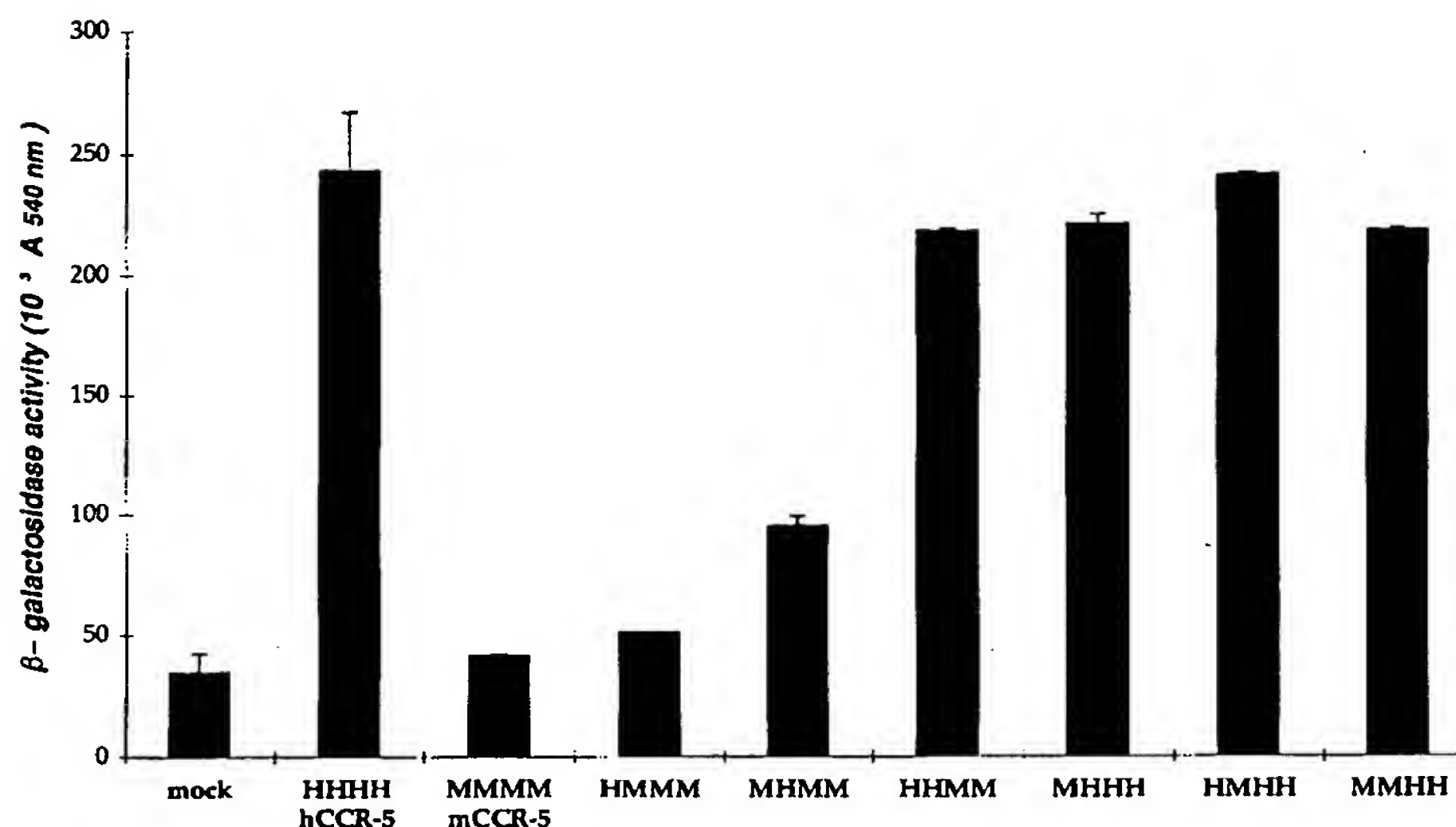


FIG. 4. Fusion assay between HIV-1<sub>ADA</sub> envelope-expressing cells and HeLa-P4 cells expressing different hCCR-5-mCCR-5 chimeras. HeLa-P4 cells transiently transfected with the different constructs were cocultivated for 24 h with HeLa-Env<sub>ADA</sub> cells in 96-well plates, lysed, and tested for  $\beta$ -galactosidase activity by a colorimetric CPRG assay as described in Materials and Methods. The histogram presented here is representative of three independent experiments done in duplicate. The results are expressed as the means of  $A_{540}$  values  $\pm$  standard deviations.

tional murine background. The chimera HHMM (for MMHH, see above) worked as well as the hCCR-5 parental molecule for ADA-induced fusion (Fig. 3 and 4) and better than the first two individual extracellular domains alone (see HMMM and MHMM). This construct confirms that the first two extracellular domains are important for coreceptor activity. Furthermore, the functionality of these two complementary chimeras (HHMM and MMHH) for HIV-1 fusion (Fig. 3 and 4) and entry (Table 1; see below) indicates that none of the intracellular and transmembrane domains of mCCR-5 are detrimental for coreceptor activity when they are placed in a functional background. This analysis shows that a combination of any single extracellular domain of hCCR-5 placed in the nonfunctional mCCR-5 background is able to mediate HIV-1<sub>ADA</sub> Env-induced fusion.

We then determined whether any domain of hCCR-5 was indispensable for the fusion process to occur. For that purpose,

we sequentially replaced segments containing each extracellular domain of hCCR-5 by an mCCR-5 segment. The resulting chimeras, MHHH, HMHH, and HHMM (for the last, see above), behaved like hCCR-5 in ADA Env-induced fusion (Fig. 4), demonstrating that none of hCCR-5 sequences were essential for coreceptor function. The alternative explanation is that mCCR-5 sequences placed in a functional hCCR-5 background adopt a different conformation and can mediate some interaction with the fusion complex. We therefore conclude that multiple domains of hCCR-5 are involved in HIV-1<sub>ADA</sub> Env-induced fusion. It was not possible to identify a single domain necessary and sufficient for fusion and likely to contain a binding site for the HIV-1 envelope.

**Different HIV-1 isolates have distinct requirements for utilization of the CCR-5 coreceptor.** Since different HIV-1 isolates can use the same coreceptor in distinct ways (66, 71), we extended our analysis to different primary SI and NSI isolates

TABLE 1. Infectivity of and syncytium formation by primary and TCLA HIV-1 isolates in cells transfected with different CCR-5 derivatives<sup>a</sup>

Viral isolate	No. of blue-stained syncytia per well <sup>b</sup> with transfected coreceptor:									
	None <sup>c</sup>	CXCR-4	hCCR-5	mCCR-5	HMMM	MHMM	HHMM	MHHH	HMHH	MMHH
<b>NSI M-tropic<sup>d</sup></b>										
ADA	1	6	965	0	146	62	643	985	1,104	976
SF-162	0	0	848	0	161	302	420	904	991	867
M23	0	0	118	0	0	8	25	124	101	4
E80	0	0	92	0	0	3	30	128	123	0
<b>SI dual-tropic</b>										
89.6	0	662	158	0	5	1	1	6	21	2
GUN-1 <sup>e</sup>	0	426	122	1	0	0	14	0	122	1

<sup>a</sup> Parental and chimeric receptors were transfected into CCC-CD4 cells, infected, and immunostained 3 days postinfection as described in Materials and Methods. Results of one of three independent experiments giving essentially the same results are presented.

<sup>b</sup> In 48-well plates.

<sup>c</sup> Transfection was with the empty pcDNA3 vector.

<sup>d</sup> SI status and tropism were previously determined (78).

<sup>e</sup> TCLA strain.



that use CCR-5 as a coreceptor and tested them in cell-free infection.

The panel of primary isolates used here have been previously described for their tropism, SI status, and coreceptor usage (78). The different hCCR-5-mCCR-5 chimeras were transiently transfected into nonpermissive feline CCC-CD4 cells before being challenged with cell-free viral supernatant. The cells were immunostained for intracellular p24 production 3 days postinfection. The results are summarized in Table 1. The ADA and SF-162 isolates were able to infect CCC-CD4 cells expressing each of the hCCR-5-mCCR-5 chimeric receptors tested, although at lower levels for HMMM and MHMM. These results parallel those obtained in the syncytium formation assay. Two other primary NSI M-tropic strains (M23 and E80) able to use the MHHH, HHMM, and HMHH chimeras were, however, unable to use HMMM, MHMM, and MMHH for efficient entry into transiently transfected CCC-CD4 cells.

Surprisingly, infection by the SI dual-tropic strains 89.6 and GUN-1 was more severely affected by modification of hCCR-5. Infection by 89.6 was efficient only in cells expressing hCCR-5 and CXCR-4, while GUN-1 was able to use only the HMHH chimeric receptor. These results were unexpected, since these isolates are able to interact with at least two divergently related chemokine receptors (CXCR-4 and CCR-5). The reason for this dependence on CCR-5 structure and sequence integrity is not yet clear but may relate to a particular envelope conformation and/or sequence of these dual-tropic isolates. The TCLA strain LAI is able to use only CXCR-4 as a coreceptor and did not use any of the chimeric CCR-5 receptors (data not shown). Since the ADA and SF-162 envelope glycoproteins were able to use all of the chimeras transfected into CCC-CD4, we assume that each receptor is expressed, folded, and transported correctly in this cell type as well and that the results observed are not cell type specific. However, variation in levels of surface expression of chimeras (particularly for HMMM) may account for some of the lower efficiency of entry observed for some isolates.

Our results show that coreceptor utilization is probably conformationally complex, involving multiple extracellular domains of CCR-5. Furthermore, different isolates are more sensitive than others to change in the sequence and/or conformation of the coreceptor. It is noteworthy that these requirements correlate with tropism, with SI dual-tropic isolates being more dependent on hCCR-5 integrity than NSI M-tropic isolates.

## DISCUSSION

In this study, we were interested in trying to determine the extracellular domains of CCR-5 important for HIV-1 entry and fusion. For this purpose we used the murine homolog of CCR-5 that has recently been cloned (10, 57) and shown to be nonfunctional as a coreceptor for HIV-1 (6, 8). This molecule has 83% identity at the amino acid level with its human counterparts (Fig. 1) but is still not able to be used as a coreceptor by a range of primary M-tropic and dual-tropic HIV-1 isolates (Fig. 3 and 4 and Table 1). Despite their high level of similarity (91%), hCCR-5 and mCCR-5 have different activities in HIV-1 coreceptor function. A rhesus macaque homolog of CCR-5, which differs from hCCR-5 at eight amino acid positions, is able to function as an HIV-1 coreceptor (16, 54). Thus, the coreceptor activities of these nonhuman CCR-5 homologs in vitro correlates with their activities in vivo: CD4<sup>+</sup> mouse lymphocytes and CD4 transgenic mice are not infected by HIV-1, whereas macaque PBMCs and macrophages are infected in vitro and in vivo by a simian-human immunodeficiency virus

recombinant that expresses the HIV-1<sub>SF-162</sub> M-tropic envelope glycoproteins (52).

Some of the changes in mCCR-5 are located in extracellular domains, suggesting that these differences may influence its function as an HIV-1 coreceptor. Subtle changes can have a dramatic effect on ligand interaction as exemplified by the chemokine binding profile of mCCR-5, which is unable to bind efficiently human MIP-1 $\alpha$  but binds human MIP-1 $\beta$  and human RANTES at levels of affinity similar to those of the corresponding murine chemokines (57). Using six hCCR-5 and mCCR-5 chimeric receptors and testing them for HIV-1<sub>ADA</sub> Env-induced cell fusion and cell-free infection with six different isolates, we showed that (i) several if not all extracellular regions of CCR-5 are implicated in the interaction with HIV-1 and (ii) there are different CCR-5 requirements for different isolates.

Using the ADA envelope glycoproteins in a syncytium formation assay, we showed that any combination of hCCR-5 segments placed in nonfunctional mCCR-5 was able to restore coreceptor function. Partial fusion was observed when we introduced a fragment spanning the human amino-terminal or E1 domain into mCCR-5. Unfortunately, reduced surface expression of HMMM may explain inefficient fusion and entry for this construct. Replacing fragments containing each extracellular domain of hCCR-5 with fragments containing the corresponding murine domain showed that none of the human domains was indispensable for the function. No single extracellular domain proved to be necessary or sufficient for CCR-5 to be used as an HIV-1 coreceptor. It was, however, not possible by this strategy to test the role of the E3 region, since this loop is identical between hCCR-5 and mCCR-5. Moreover, no important roles were noted for transmembrane and intracellular regions in HIV-1 coreceptor activity with the chimeras used here. The implication from these data is that the interaction between Env and CCR-5 involves more than one region of the coreceptor. Clearly, mCCR-5 is incapable of making some key interactions resulting in lack of fusion activity, but this deficiency can be repaired in several ways.

Using a panel of primary NSI M-tropic and SI dual-tropic isolates, we found that the requirements were not the same between these phenotypically different strains. Cell-free virus infection by the ADA and SF-162 isolates gave similar results in the syncytium formation assay. However, two other primary NSI M-tropic strains (M23 and E80) were able to fuse only with some of the chimeric coreceptors. For these two isolates, we were unable to challenge transfected cells with as high an input of virus as that used for ADA and SF-162. Thus, the lower levels of or lack of infection by these two viruses noted for some of the constructs may partly be due to the low viral input. In the case of MMHH, however, it is clear that while full infectivity by ADA and SF-162 was observed, M23 and E80 plating was at background levels. Further studies will be needed to assess the extent of variation in the use of CCR-5 between distinct NSI M-tropic strains. Bieniasz et al. recently showed a differential usage of CCR-5 between the two NSI M-tropic isolates ADA and Ba-L in cell-cell fusion that correlated to the V3 loop sequence (8).

In contrast, the SI dual-tropic strains generally could use few of the chimeric receptors. The strain specificity reported here was not totally unexpected, as it has been shown that different HIV-1 strains have different requirements for interaction with the N-terminal extracellular domain of CXCR-4 (66) or CCR-5 (71). A recent study using CCR-5-CCR-2b chimeras also showed that HIV-1 must interact with multiple domains, the N-terminal and E1 domains being the most important (71). Using the 89.6 dual-tropic isolates, Rucker et al. showed that

this isolate has requirements different from those of JR-FL for CCR-5 use. However, the interpretation of the results was complicated by the fact that the partner used for the chimeras was CCR-2b, which is also used as a coreceptor by 89.6. We showed here that the dual-tropic 89.6 and GUN-1 isolates were inefficient in using most of the hCCR-5-mCCR-5 chimeras for entry, despite the fact that the level of infection of CCR-5-expressing cells was the same as for M-tropic isolates.

The differences observed between NSI M-tropic and SI dual-tropic isolates for CCR-5 usage are surprising, since the latter isolates can use multiple coreceptors to gain entry into cells: 89.6 and GUN-1 can use CCR-5 and CXCR-4 (33, 78), and in addition, 89.6 can use CCR-3 and CCR-2b (34). Thus, they are able to use different divergent members of the chemokine receptor family for entry into cells and yet subtle changes in CCR-5 abolish their entry. In contrast, M-tropic isolates proved less sensitive to changes in CCR-5 despite the fact that their coreceptor usage is more restricted. This apparent paradox suggests that multiple ways of interacting with CCR-5 exist among HIV-1 strains, which may influence their ability to gain the use of other coreceptors. It is possible that evolution of a CCR-5-specific HIV-1 envelope to exploit more than one coreceptor may lower its affinity for CCR-5, making it more sensitive to changes in the sequence. The changes that take place in NSI isolate envelope glycoprotein sequences *in vivo* during the NSI-to-SI phenotypic switch (14, 18, 39, 49) must reflect the ability of their envelope glycoproteins to interact with CXCR-4 (46). These changes may not be compatible with interaction with CCR-5 at the same sites or with the same affinity. Alternatively, the ability of gp120 to interact with the coreceptor is necessary but not sufficient for the entry process to be completed.

It has recently been shown that sequences inside the V3 loop known to influence tropism (17, 21, 44, 56, 61, 75, 80, 85, 87) and post-CD4 binding events (7, 40, 45, 64) also influence coreceptor usage (22, 33, 67), formation of the CD4-CCR-5-gp120 ternary complex (81, 88), and sensitivity to  $\beta$ -chemokine-mediated inhibition of entry (27, 46). The V3 loop is therefore a candidate region for interaction with CCR-5 but is probably not the only Env region involved. In fact, it has been shown that a recombinant NSI M-tropic isolate containing a TCLA strain V3 loop was still able to infect PBMCs and macrophages. This recombinant probably still interacted with CCR-5 despite the TCLA strain-derived V3 loop sequence (15, 62). Other variable loops like V1 and V2 that influence tropism and the SI ability of HIV-1 (11, 47, 79) may also be involved. Other factors, such as a possible CD4-CCR-5 interaction, must also be taken into account to explain the precise mechanism of interaction of the HIV-1 envelope glycoproteins with the coreceptor. Results from a study by Wu et al. (88) showing that a soluble form of CD4 consisting of the two amino-terminal domains (D1D2 sCD4) can inhibit MIP-1 $\alpha$  and MIP-1 $\beta$  binding to CCR-5 in the absence of gp120 support this possibility. Furthermore, differences in affinities for the binding receptor CD4 may also influence the way different HIV-1 strains interact with the receptor complex (48).

During the course of this work, two studies also reported by a similar approach that several CCR-5 extracellular domains were involved in HIV-1 entry (6, 8). In contrast to our results, Atchinson et al. (6) showed that an MMHH construct was only minimally active for Ba-L, while our equivalent construct fully supported entry of two NSI M-tropic isolates: ADA and SF-162. Bieniasz et al. using an envelope-mediated cell fusion assay reached conclusions similar to those of our study: they observed differential usage of hCCR-5 and mCCR-5 chimeras between distinct NSI M-tropic isolates (ADA and Ba-L) and

restrictive utilization of the chimeric receptors by the 89.6 isolate. Our analysis further shows that this is also true in cell-free virus infection and that it is likely to be a general feature of dual-tropic strains to interact less efficiently or differently with CCR-5.

In summary, we have shown that multiple domains of CCR-5 are involved in HIV-1 entry and syncytium formation. Six different CCR-5-using isolates were tested and shown to vary in coreceptor requirements for cell entry. These distinct coreceptor needs correlated with tropism, with dual-tropic strains being more sensitive to changes in their CCR-5 sequences. It will be interesting to see how these dual-tropic isolates interact with the other major coreceptor, CXCR-4. Other chimeras and mutants of CCR-5 and CXCR-4 need to be tested to precisely understand the HIV-1 entry mechanism.

#### ACKNOWLEDGMENTS

We thank Ron Collman (University of Pennsylvania, Philadelphia), Pat Gray (ICOS Corp.), and Anne Brelot and Marc Alizon (Institut Cochin de Génétique Moléculaire) for providing reagents; Paul Kellam for help with sequencing; and Matthias Dittmar and Tim Wells for helpful discussions and critical readings of the manuscript.

This work was supported by the Medical Research Council. L.P. was supported by an MRC-INSERM exchange fellowship.

#### REFERENCES

- Adachi, A., H. E. Gendelmann, S. Koenig, T. Folks, R. Willey, A. Rabson, and M. A. Martin. 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and non-human cells transfected with an infectious molecular clone. *J. Virol.* 59:284-291.
- Alkhatib, G., C. C. Broder, and E. A. Berger. 1996. Cell type-specific fusion cofactors determine human immunodeficiency virus type 1 tropism for T-cell lines versus primary macrophages. *J. Virol.* 70:5487-5494.
- Alkhatib, G., C. Combadiere, C. C. Broder, Y. Feng, P. E. Kennedy, P. M. Murphy, and A. E. Berger. 1996. CC CKR5: a RANTES, MIP-1 $\alpha$ , MIP-1 $\beta$  receptor as fusion cofactor for macrophage-tropic HIV-1. *Science* 272:1955-1958.
- Arenzana-Seisdedos, F., J.-L. Virelizier, D. Rousset, I. Clark-Lewis, P. Loetscher, B. Moser, and M. Baggiolini. 1996. HIV blocked by chemokine antagonist. *Nature* 383:400.
- Ashorn, P. A., E. A. Berger, and B. Moss. 1990. Human immunodeficiency virus envelope glycoprotein/CD4-mediated fusion of nonprimate cells with human cells. *J. Virol.* 64:2149-2156.
- Atchinson, R. E., J. Gosling, F. S. Monteclaro, C. Franci, L. Digilio, I. F. Charo, and M. A. Goldsmith. 1996. Multiple extracellular elements of CCR5 and HIV-1 entry: dissociation from response to chemokines. *Science* 274:1924-1926.
- Bergeron, L., N. Sullivan, and J. Sodroski. 1992. Target cell-specific determinants of membrane fusion within the human immunodeficiency virus type 1 gp120 third variable domain and gp41 amino terminus. *J. Virol.* 66:2389-2397.
- Bieniasz, P. D., R. A. Fridell, I. Aramori, S. S. G. Ferguson, M. G. Caron, and B. R. Cullen. HIV-1 induced cell fusion is mediated by multiple regions within both the viral envelope and the CCR-5 co-receptor. *EMBO J.*, in press.
- Bleul, C. C., M. Farzan, H. Choe, C. Parolin, I. Clark-Lewis, J. Sodroski, and T. A. Springer. 1996. The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry. *Nature* 382:829-833.
- Boring, L., J. Gosling, F. S. Monteclaro, A. J. Lusis, C.-L. Tsou, and I. F. Charo. 1996. Molecular cloning and functional expression of murine JE (monocyte chemoattractant protein 1) and murine macrophage inflammatory protein 1 $\alpha$  receptors. *J. Biol. Chem.* 271:7551-7558.
- Boyd, M. T., G. R. Simpson, A. J. Cann, M. A. Johnson, and R. A. Weiss. 1993. A single amino acid substitution in the V1 loop of human immunodeficiency virus type 1 gp120 alters cellular tropism. *J. Virol.* 67:3649-3652.
- Broder, C., and E. A. Berger. 1995. Fusogenic selectivity of the envelope glycoprotein is a major determinant of human immunodeficiency virus type 1 tropism for CD4<sup>+</sup> T-cell lines vs. primary macrophages. *Proc. Natl. Acad. Sci. USA* 92:9004-9008.
- Broder, C. C., D. S. Dimitrov, R. Blumenthal, and E. A. Berger. 1993. The block to HIV-1 envelope glycoprotein-mediated membrane fusion in animal cells expressing human CD4 can be overcome by a human cell component(s). *Virology* 193:483-491.
- Callahan, L. 1994. HIV-1 virion-cell interactions: an electrostatic model of pathogenicity and syncytium formation. *AIDS Res. Hum. Retroviruses* 10:231-233.



15. Carrillo, A., and L. Ratner. 1996. Human immunodeficiency virus type 1 tropism for T-lymphoid cell lines: role of the V3 loop and C4 envelope determinants. *J. Virol.* 70:1301-1309.
16. Chen, Z., P. Zhou, D. D. Ho, N. R. Landau, and P. A. Marx. 1997. Genetically divergent strains of simian immunodeficiency virus use CCR5 as a coreceptor for entry. *J. Virol.* 71:2705-2714.
17. Cheng-Mayer, C., M. Oulroga, J. W. Tung, D. Dina, and J. A. Levy. 1990. Virol determinants of human immunodeficiency virus type 1 T-cell or macrophage tropism, cytopathogenicity, and CD4 antigen modulation. *J. Virol.* 64:4390-4398.
18. Cheng-Mayer, C., T. Shioda, and J. A. Levy. 1991. Host range, replicative, and cytopathic properties of human immunodeficiency virus type 1 are determined by very few amino acid changes in *tat* and gp120. *J. Virol.* 65:6931-6941.
19. Cheng-Mayer, C., C. Weiss, D. Seto, and J. A. Levy. 1989. Isolates of human immunodeficiency virus type 1 from the brain may constitute a special group of the AIDS virus. *Proc. Natl. Acad. Sci. USA* 86:8575-8579.
20. Chesebro, B., R. Buller, J. Portis, and K. Wehrly. 1990. Failure of human immunodeficiency virus entry and infection in CD4-positive human brain and skin cells. *J. Virol.* 64:215-221.
21. Chesebro, B., K. Wehrly, J. Nishio, and S. Perryman. 1992. Macrophage-tropic human immunodeficiency virus isolates from different patients exhibit unusual V3 envelope sequence homogeneity in comparison with T-cell tropic isolates: definition of critical amino acids involved in cell tropism. *J. Virol.* 66:6547-6554.
22. Choe, H., M. Farzan, Y. Sun, N. Sullivan, B. Rollins, P. D. Ponath, L. Wu, C. R. Mackay, G. LaRosa, W. Newman, N. Gerard, C. Gerard, and J. Sodroski. 1996. The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell* 85:1135-1148.
23. Clapham, P. R., D. Blanc, and R. A. Weiss. 1991. Specific cell surface requirements for the infection of CD4-positive cells by human immunodeficiency virus types 1 and 2 and by simian immunodeficiency virus. *Virology* 181:703-715.
24. Clapham, P. R., A. McKnight, and R. A. Weiss. 1992. Human immunodeficiency virus type 2 infection and fusion of CD4-negative human cell lines: induction and enhancement by soluble CD4. *J. Virol.* 66:3531-3537.
25. Clavel, F., and P. Charneau. 1994. Fusion from without directed by human immunodeficiency virus particles. *J. Virol.* 68:1179-1185.
26. Cocchi, F., A. L. DeVico, A. Garzino-Demo, S. K. Arya, R. C. Gallo, and P. Lusso. 1995. Identification of RANTES, MIP-1 $\alpha$ , and MIP-1 $\beta$  as the major HIV-suppressive factors produced by CD8<sup>+</sup> T cells. *Science* 270:1811-1815.
27. Cocchi, F., A. L. DeVico, A. Garzino-Demo, A. Cara, R. C. Gallo, and P. Lusso. 1996. The V3 domain of the gp120 envelope glycoprotein is critical for chemokine-mediated blockade of infection. *Nat. Med.* 2:1244-1247.
28. Collman, R., J. W. Balliet, S. A. Gregory, H. Friedman, D. L. Kolson, N. Nathanson, and A. Srinivasan. 1992. An infectious molecular clone of an unusual macrophage-tropic and highly cytopathic strain of human immunodeficiency virus type 1. *J. Virol.* 66:7517-7521.
29. Combadiere, C., S. K. Ahuja, H. Lee Tiffany, and P. M. Murphy. 1996. Cloning and functional expression of CC CKR5, a human monocyte CC chemokine receptor selective for MIP-1 $\alpha$ , MIP-1 $\beta$ , and RANTES. *J. Leukocyte Biol.* 60:147-152.
30. Connor, R. I., K. E. Sheridan, D. Ceradini, S. Choe, and N. R. Landau. 1997. Change in coreceptor use correlates with disease progression in HIV-1-infected individuals. *J. Exp. Med.* 185:621-628.
31. Dean, M., M. Carrington, C. Winkler, G. A. Huttley, M. W. Smith, R. Allikmets, J. J. Goedert, S. P. Buchbinder, E. Vittinghoff, E. Gomperts, S. Donfield, D. Vlahov, R. Kaslow, A. Saah, C. Rinaldo, R. Detels, Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study, and S. J. O'Brien. 1996. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the *CCR5* structural gene. *Science* 273:1856-1862.
32. Deng, H., R. Liu, W. Elmeier, S. Choe, D. Unutmaz, M. Burkhart, P. Di Marzio, S. Marmon, R. E. Sutton, C. M. Hill, C. B. Davis, S. C. Peiper, T. J. Schall, D. R. Littman, and N. R. Landau. 1996. Identification of a major co-receptor for primary isolates of HIV-1. *Nature* 381:661-666.
33. Dittmar, M. T., A. McKnight, G. Simmons, P. R. Clapham, R. A. Weiss, and P. Simmonds. 1997. HIV-1 tropism and co-receptor usage. *Nature* 385:495-496.
34. Doranz, B. J., J. Rucker, Y. Yi, R. J. Smyth, M. Samson, S. C. Peiper, M. Parmentier, R. G. Collman, and R. W. Doms. 1996. A dual-tropic primary HIV-1 isolate that uses fusin and the beta-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors. *Cell* 85:1149-1158.
35. Dragic, T., P. Charneau, F. Clavel, and M. Alizon. 1992. Complementation of murine cells for human immunodeficiency virus envelope/CD4-mediated fusion in human-murine heterokaryons. *J. Virol.* 66:4794-4802.
36. Dragic, T., V. Litwin, G. Allaway, S. R. Martin, Y. Huang, K. A. Nagashima, C. Cayanan, P. J. Maddon, R. A. Koup, J. P. Moore, and W. A. Paxton. 1996. HIV-1 entry into CD4<sup>+</sup> cells is mediated by the chemokine receptor CC-CKR5. *Nature* 381:667-673.
37. Evan, G. I., G. K. Lewis, G. Ramsay, and J. M. Bishop. 1985. Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. *Mol. Cell. Biol.* 5:3610-3616.
38. Feng, Y., C. C. Broder, P. E. Kennedy, and E. A. Berger. 1996. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane G-protein coupled receptor. *Science* 272:872-877.
39. Fouchier, R. A. M., M. Groenik, N. A. Kootstra, M. Tersmette, H. G. Huismann, F. Miedema, and H. Schuitemaker. 1992. Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. *J. Virol.* 66:3183-3187.
40. Freed, E. O., D. J. Myers, and R. Risser. 1991. Identification of the principal neutralization determinant of human immunodeficiency virus type 1 as a fusion domain. *J. Virol.* 65:190-194.
41. Harrington, R., and A. P. Geballe. 1993. Cofactor requirements for human immunodeficiency virus type 1 entry into a CD4-expressing human cell line. *J. Virol.* 67:5939-5947.
42. He, J., Y. Chen, M. Farzan, H. Choe, A. Ohagen, S. Gartner, J. Busciglio, X. Yang, W. Hofmann, W. Newman, C. R. Mackay, J. Sodroski, and D. Gabuzda. 1997. CCR3 and CCR5 are co-receptors for HIV-1 infection of microglia. *Nature* 385:645-649.
43. Huang, Y., W. A. Paxton, S. M. Wolinsky, A. U. Neumann, L. Zhang, T. He, S. Kang, D. Ceradini, Z. Jin, K. Yazdanbakhsh, K. Kunstman, D. Erickson, E. Dragon, N. R. Landau, J. Phair, D. D. Ho, and R. A. Koup. 1996. The role of a mutant CCR5 allele in HIV-1 transmission and disease progression. *Nat. Med.* 2:1240-1243.
44. Hwang, S. R., T. J. Boyle, K. Lyerly, and B. R. Cullen. 1992. Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1. *Science* 253:71-74.
45. Ivanoff, L. A., J. W. Dubay, J. F. Morris, S. J. Roberts, L. Gutshall, E. J. Sternberg, E. Hunter, T. J. Matthews, and S. R. Petteway, Jr. 1992. V3 loop of the HIV-1 envelope protein is essential for virus infectivity. *Virology* 187:423-432.
46. Jansson, M., M. Popovic, A. Karlsson, F. Cocchi, P. Rossi, J. Albert, and H. Wigzell. 1996. Sensitivity to inhibition by  $\beta$ -chemokines correlates with biological phenotypes of primary HIV-1 isolates. *Proc. Natl. Acad. Sci. USA* 93:15382-15387.
47. Koito, A., G. Harrowe, J. A. Levy, and C. Cheng-Mayer. 1994. Functional role of the V1/V2 region of human immunodeficiency virus type 1 envelope glycoprotein gp120 in infection of primary macrophages and soluble CD4 neutralization. *J. Virol.* 68:2253-2259.
48. Kozak, S. L., E. J. Platt, N. Madani, F. E. Ferro, Jr., K. Peden, and D. Kabat. 1997. CD4, CXCR-4, and CCR-5 dependencies for infections by primary patient and laboratory-adapted isolates of human immunodeficiency virus type 1. *J. Virol.* 71:873-882.
49. Kuiken, C. L., J. J. de Jong, E. Baan, W. Keulen, M. Tersmette, and J. Goudsmith. 1992. Evolution of V3 envelope domain in proviral sequences and isolates of human immunodeficiency virus type 1 during the transition of the viral biological phenotype. *J. Virol.* 66:4622-4627.
50. Lapham, C. K., J. Ouyang, B. Chandrasekhar, N. Y. Nguyen, D. S. Dimitrov, and H. Golding. 1996. Evidence for cell-surface association between fusin and the CD4-gp120 complex in human cell lines. *Science* 274:602-605.
51. Liu, R., W. A. Paxton, S. Choe, D. Ceradini, S. R. Martin, R. Horuk, M. E. MacDonald, H. Stuhlmann, R. A. Koup, and N. R. Landau. 1996. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 86:367-377.
52. Luciw, P. A., E. Pratt-Lowe, K. E. S. Shaw, J. A. Levy, and C. Cheng-Mayer. 1995. Persistent infection of rhesus macaques with T-cell-line-tropic and macrophage-tropic clones of simian/human immunodeficiency viruses (SHIV). *Proc. Natl. Acad. Sci. USA* 92:7490-7494.
53. Maddon, P. J., A. G. Dalgleish, J. S. McDougall, P. R. Clapham, R. A. Weiss, and R. Axel. 1986. The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell* 47:333-348.
54. Marcon, L., H. Choe, K. A. Martin, M. Farzan, P. D. Ponath, L. Wu, W. Newman, N. Gerard, C. Gerard, and J. Sodroski. 1997. Utilization of C-C chemokine receptor 5 by the envelope glycoproteins of a pathogenic simian immunodeficiency virus, SIV<sub>mac</sub>239. *J. Virol.* 71:2522-2527.
55. McKnight, A., P. R. Clapham, and R. A. Weiss. 1994. HIV-2 and SIV infection of nonprimate cell lines expressing human CD4: restrictions to replication at distinct stages. *Virology* 201:8-18.
56. McKnight, A., R. A. Weiss, C. Shotton, Y. Takeuchi, H. Hoshino, and P. R. Clapham. 1995. Change in tropism upon immune escape by human immunodeficiency virus. *J. Virol.* 69:3167-3170.
57. Meyer, A., A. J. Coyle, A. E. I. Proudfoot, T. N. C. Wells, and C. A. Power. 1996. Cloning and characterization of a novel murine macrophage inflammatory protein-1 $\alpha$  receptor. *J. Biol. Chem.* 271:14445-14451.
58. Moore, J. P., B. A. Jameson, R. A. Weiss, and Q. J. Sattentau. 1993. The HIV-cell fusion reaction, p. 233-289. *In* J. Bentz (ed.), *Viral fusion mechanisms*. CRC Press, Boca Raton, Fla.
59. Nussbaum, O., C. C. Broder, and E. A. Berger. 1994. Fusogenic mechanisms of enveloped-virus glycoproteins analyzed by a novel recombinant vaccinia virus-based assay quantitating cell fusion-dependent reporter gene activation. *J. Virol.* 68:5411-5422.
60. Oberlin, E., A. Amara, F. Bachelier, C. Bessia, J.-L. Virelizier, F. Arenzana-



- Seisdedos, O., Schwartz, J.-M., Heard, I., Clark-Lewis, D. F., Legler, M., Loetscher, M., Baggolini, and B. Moser. 1996. The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. *Nature* 382:833-835.
61. O'Brien, W. A., Y. Koyanagi, A. Namazie, J.-Q. Zhao, A. Diagne, K. Idler, J. A. Zack, and I. S. Y. Chen. 1990. HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 outside the CD4-binding domain. *Nature* 348:69-73.
  62. O'Brien, W. A., M. Sumner-Smith, S.-H. Mao, S. Sadeghi, J.-Q. Zhao, and I. S. Y. Chen. 1996. Anti-human immunodeficiency virus type 1 activity of an oligocationic compound mediated via gp120 V3 interactions. *J. Virol.* 70:2825-2831.
  63. Oravec, T., M. Pall, and M. A. Norcross. 1996.  $\beta$ -Chemokine inhibition of monocyctotropic HIV-1 infection. Interference with a postbinding fusion step. *J. Immunol.* 157:1329-1332.
  64. Page, K. A., S. M. Stearns, and D. R. Littman. 1992. Analysis of mutations in the V3 domain of gp160 that affect fusion and infectivity. *J. Virol.* 66:524-533.
  65. Paxton, W. A., S. R. Martin, D. Tse, T. R. O'Brien, J. Skurnick, N. L. VanDevanter, N. Padian, J. F. Braun, D. P. Kotler, S. M. Wolinsky, and R. A. Koup. 1996. Relative resistance to HIV-1 infection of CD4 lymphocytes from persons who remain uninfected despite multiple high-risk sexual exposure. *Nat. Med.* 2:412-417.
  66. Picard, L., D. A. Wilkinson, A. McKnight, P. W. Gray, J. A. Hoxie, P. R. Clapham, and R. A. Weiss. Role of the amino-terminal extracellular domain of CXCR-4 in human immunodeficiency virus type 1 entry. *Virology*, in press.
  67. Pleskoff, O., N. Sol, B. Labrousse, and M. Alizon. 1997. Human immunodeficiency virus strains differ in their ability to infect CD4<sup>+</sup> cells expressing the rat homolog of CXCR-4 (fusin). *J. Virol.* 71:3259-3262.
  68. Pleskoff, O., C. Treboute, A. Brelot, N. Heveker, M. Seman, and M. Alizon. A chemokine receptor encoded by the human cytomegalovirus is a cofactor for HIV-1 entry. *Science*, in press.
  69. Premack, B. A., and T. J. Schall. 1996. Chemokine receptors: gateways to inflammation and infection. *Nat. Med.* 2:1174-1178.
  70. Raport, C. J., J. Gosling, V. L. Schweickart, P. W. Gray, and I. F. Charo. 1996. Molecular cloning and functional characterization of a novel human CC chemokine receptor (CCR5) for RANTES, MIP- $\beta$ , and MIP-1 $\alpha$ . *J. Biol. Chem.* 271:17161-17166.
  71. Rucker, J., M. Samson, B. J. Doranz, F. Libert, J. F. Bereson, Y. Yi, R. G. Collman, C. C. Broder, G. Vassart, R. W. Doms, and M. Parmentier. 1996. Regions in  $\beta$ -chemokine receptors CCR5 and CCR2b that determine HIV-1 cofactor specificity. *Cell* 87:437-446.
  72. Samson, M., O. Labbe, C. Mollereau, G. Vassart, and M. Parmentier. 1996. Molecular cloning and functional expression of a new human CC-chemokine receptor gene. *Biochemistry* 11:3362-3367.
  73. Samson, M., F. Libert, B. J. Doranz, J. Rucker, C. Liesnard, C.-M. Farber, S. Saragosti, C. Lapoumeroulie, J. Cogniaux, C. Forcellie, G. Muyldermans, C. Verhofstede, G. Burtonboy, M. Georges, T. Imai, S. Rana, Y. Yi, R. J. Smyth, R. G. Collman, R. W. Doms, G. Vassart, and M. Parmentier. 1996. Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* 382:722-725.
  74. Schwartz, O., M. Alizon, J. M. Heard, and O. Danos. 1994. Impairment of T cell receptor-dependent stimulation in CD4<sup>+</sup> lymphocytes after contact with membrane-bound HIV-1 envelope glycoprotein. *Virology* 198:360-365.
  75. Shioda, T., J. A. Levy, and C. Cheng-Mayer. 1991. Macrophage and T-cell line tropisms of HIV-1 are determined by specific regions of the envelope gp120 gene. *Nature* 349:167-169.
  76. Simmons, G., P. R. Clapham, L. Picard, R. E. Offord, M. M. Rosenkilde, T. W. Schwartz, R. Buser, T. N. C. Wells, and A. E. I. Proudfoot. 1997. Potent inhibition of HIV-1 infectivity in macrophages and lymphocytes by a novel CCR5 antagonist. *Science* 276:276-279.
  77. Simmons, G., A. McKnight, Y. Takeuchi, H. Hoshino, and P. R. Clapham. 1995. Cell-to-cell fusion, but not virus entry in macrophages by T-cell line tropic HIV-1 strains: a V3 loop-determined restriction. *Virology* 209:696-700.
  78. Simmons, G., D. Wilkinson, J. D. Reeves, M. T. Dittmar, S. Beddows, J. Weber, G. Carnegie, U. Desselberger, P. W. Gray, R. A. Weiss, and P. R. Clapham. 1996. Primary, syncytium-inducing human immunodeficiency virus type 1 isolates are dual-tropic and most can use either Lestr or CCR5 as coreceptors for virus entry. *J. Virol.* 70:8355-8360.
  79. Sullivan, N., M. Thali, C. Furman, D. D. Ho, and J. Sodroski. 1993. Effects of amino acid changes in the V2 region of the human immunodeficiency virus type 1 gp120 glycoprotein on subunit association, syncytium formation, and recognition by neutralizing antibody. *J. Virol.* 67:3674-3679.
  80. Takeuchi, Y., M. Akutsu, K. Murayama, N. Shimizu, and H. Hoshino. 1991. Host-range mutant of human immunodeficiency virus type 1: modification of cell tropism by a single point mutation at the neutralization epitope in the *env* gene. *J. Virol.* 65:1710-1718.
  81. Trkola, A., T. Dragic, J. Arthos, J. M. Binley, W. C. Olson, G. P. Allaway, C. Cheng-Mayer, J. Robinson, P. J. Maddon, and J. P. Moore. 1996. CD4-dependent, antibody-sensitive interactions between HIV-1 and its co-receptor CCR-5. *Nature* 384:184-187.
  82. Wain-Hobson, S., J.-P. Vartanian, M. Henry, N. Chenciner, R. Cheynier, S. Delassus, L. P. Martins, M. Sala, M. T. Nugeyre, D. Guetard, D. Klatzmann, J.-C. Gluckmann, W. Rosenbaum, F. Barré-Sinoussi, and L. Montagnier. 1991. LAV revisited: origins of the early HIV-1 isolates from Institut Pasteur. *Science* 252:961-965.
  83. Weiss, R. A. 1993. Cellular receptors and viral glycoproteins involved in retrovirus entry, p. 1-108. *In* J. A. Levy (ed.), *The retroviridae*, vol. 2. Plenum Press, New York, N.Y.
  84. Wells, T. N., C. A. Power, M. Lusti-Narasimhan, A. J. Hoogewerf, R. M. Cooke, C. W. Chung, M. C. Peitsch, and A. E. Proudfoot. 1996. Selectivity and antagonism of chemokine receptors. *J. Leukocyte Biol.* 59:53-60.
  85. Westervelt, P., H. E. Gendelman, and L. Ratner. 1991. Identification of a determinant within the human immunodeficiency virus type 1 surface envelope glycoprotein critical for productive infection of primary monocytes. *Proc. Natl. Acad. Sci. USA* 88:3097-3101.
  86. Westervelt, P., D. Trowbridge, L. Epstein, B. Blumberg, Y. Li, B. Hahn, G. Shaw, R. Price, and L. Ratner. 1992. Macrophage tropism determinants of human immunodeficiency virus type 1 in vivo. *J. Virol.* 66:2577-2582.
  87. Willey, R. L., T. S. Theodore, and M. A. Martin. 1994. Amino acid substitutions in the human immunodeficiency virus type 1 gp120 V3 loop that change tropism also alter physical properties of the virion envelope. *J. Virol.* 68:4409-4419.
  88. Wu, L., N. P. Gerard, R. Wyatt, H. Choe, C. Parolin, N. Ruffing, A. Borsetti, A. A. Cardoso, E. Desjardin, W. Newman, C. Gerard, and J. Sodroski. 1996. CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. *Nature* 384:179-183.
  89. Zhang, L., Y. Huang, T. He, Y. Cao, and D. D. Ho. 1996. HIV-1 subtype and second-receptor use. *Nature* 383:768.

## Role of the First and Third Extracellular Domains of CXCR-4 in Human Immunodeficiency Virus Coreceptor Activity

ANNE BRELOT, NIKOLAUS HEVEKER, OLIVIER PLESKOFF, NATHALIE SOL, AND MARC ALIZON\*

*INSERM U.332, Institut Cochin de Génétique Moléculaire, 75014 Paris, France*

Received 17 January 1997/Accepted 4 March 1997

The CXCR-4 chemokine receptor and CD4 behave as coreceptors for cell line-adapted human immunodeficiency virus types 1 and 2 (HIV-1 and HIV-2) and for dual-tropic HIV strains, which also use the CCR-5 coreceptor. The cell line-adapted HIV-1 strains LAI and NDK and the dual-tropic HIV-2 strain ROD were able to infect CD4<sup>+</sup> cells expressing human CXCR-4, while only LAI was able to infect cells expressing the rat homolog of CXCR-4. This strain selectivity was addressed by using human-rat CXCR-4 chimeras. All chimeras tested mediated LAI infection, but only those containing the third extracellular domain (e3) of human CXCR-4 mediated NDK and ROD infection. The e3 domain might be required for the functional interaction of NDK and ROD, but not LAI, with CXCR-4. Alternatively, LAI might also interact with e3 but in a different way. Monoclonal antibody 12G5, raised against human CXCR-4, did not stain cells expressing rat CXCR-4. Chimeric human-rat CXCR-4 allowed us to map the 12G5 epitope in the e3 domain. The ability of 12G5 to neutralize infection by certain HIV-1 and HIV-2 strains is also consistent with the role of e3 in the coreceptor activity of CXCR-4. The deletion of most of the amino-terminal extracellular domain (e1) abolished the coreceptor activity of human CXCR-4 for ROD and NDK but not for LAI. These results indicate that HIV strains have different requirements for their interaction with CXCR-4. They also suggest differences in the interaction of dual-tropic HIV with CCR-5 and CXCR-4.

Human immunodeficiency virus types 1 and 2 (HIV-1 and HIV-2) infect cells by a membrane fusion process mediated by their envelope glycoproteins (gp120/gp41, or Env) that requires the interaction of gp120 with the cellular receptor CD4 (reviewed in references 27 and 41). The complete resistance of several nonhuman and human CD4<sup>+</sup> cell lines to HIV-1 infection and to the formation of syncytia with Env<sup>+</sup> cells suggests that other cellular factors are required for Env-mediated fusion (3, 6, 10, 13, 19, 25). The existence of such factors could also explain cell tropism differences between clinical isolates (primary strains) and cell line-adapted strains (23). Primary strains generally have a non-syncytium-inducing (NSI) phenotype and replicate in peripheral blood mononuclear cells and macrophages but not in CD4<sup>+</sup> cell lines. Certain primary strains, usually isolated at late stages of HIV infection, have a syncytium-inducing (SI) phenotype and can replicate in CD4<sup>+</sup> cell lines and peripheral blood mononuclear cells, but most HIV strains with these properties have been obtained by ex vivo adaptation. Finally, certain HIV strains, termed dual tropic, display features of both primary and cell line-adapted strains. These phenotypic differences among HIV strains are principally supported by genetic differences in their gp120 surface envelope proteins, in particular in the third hypervariable (V3) domain (reviewed in references 18, 27, and 41).

The cellular factors responsible for the permissiveness of CD4<sup>+</sup> cells to HIV-1 entry were found to be members of the chemokine receptor family. Chemokines are small (molecular weight, 8,000 to 10,000) soluble proteins that mediate leukocyte chemotaxis by interacting with a family of G-protein-coupled receptors with seven membrane-spanning (TM) domains. Chemokines are classified into two groups,  $\alpha$  (or CXC) and  $\beta$  (or CC), depending upon the position of two conserved

cysteines in their amino-terminal domains (reviewed in reference 28). A factor allowing entry of cell line-adapted HIV-1 strains was first identified by genetic complementation of murine CD4<sup>+</sup> cells and was named fusin (16). It was later shown to be a receptor for stromal cell-derived factor 1 (SDF-1), a CXC chemokine, and renamed CXCR-4 (5, 30). This chemokine receptor can also be used for cell entry by HIV-2 strains, in both CD4-dependent (33) and CD4-independent (15) modes. It does not allow infection by NSI (or macrophage-tropic) HIV-1 strains, which were found to use CCR-5, a receptor for the CC chemokines RANTES, MIP-1 $\alpha$ , and MIP-1 $\beta$  (2, 9, 11, 12, 14). Dual-tropic HIV-1 (9, 12), primary HIV-1 with an SI phenotype (37, 44), and HIV-2 strains such as ROD (38) can use both CXCR-4 and CCR-5 for cell entry. In addition to these chemokine receptors, certain HIV-1 strains seem to be able to use CCR-3 and CCR-2b (9, 12).

The mechanism by which chemokine receptors participate in the process of HIV-1 entry is not entirely known, but different experiments suggest that gp120 interacts with CXCR-4 or CCR-5 (22, 40, 43). This interaction is markedly stronger in the presence of CD4, which suggests that chemokine receptors and CD4 behave as coreceptors for HIV. The identification of the domains of the chemokine receptors that interact with gp120 is important for understanding the process of HIV-1 entry. We have previously shown that human CXCR-4 was a coreceptor for the genetically divergent HIV-1 strains LAI (clade B) and NDK (clade D) and for HIV-2<sub>ROD</sub>. In contrast, the rat homolog of CXCR-4 was able to efficiently mediate infection by LAI but not by NDK and ROD (33). This strain selectivity was used to identify domains of CXCR-4 supporting HIV coreceptor activity.

### MATERIALS AND METHODS

**Cell lines and viral strains.** The U373MG-CD4 cell line (19) and HeLa cell lines stably expressing wild-type or chimeric LAI Env (33, 36) have been described previously. HIV-1<sub>NDK</sub> (39) and HIV-1<sub>NL4.3</sub> (1) were propagated in the T-cell line CEM. HIV-1<sub>LAI</sub> was produced by transfection of HeLa cells with a

\* Corresponding author. Mailing address: INSERM U.332, Institut Cochin de Génétique Moléculaire, 22 rue Méchain, 75014 Paris, France. Phone: 33-1-40 51 64 86. Fax: 33-1-40 51 77 49. E-mail: alizon@cochin.inserm.fr.



recombinant provirus (32). HIV-2<sub>ROD</sub> was produced by a chronically infected cell line (7).

**Construction of chimeric CXCR-4.** The R<sub>c</sub>/CMV vectors allowing expression of human and rat CXCR-4 from the cytomegalovirus immediate-early promoter have been described previously (33). The coding sequence of the rat CXCR-4 cDNA (kindly provided by R. S. Duman) was determined by direct automated sequencing of the expression vector. Several differences from the previously published sequence (42) (GenBank accession no. U54791) were noted, all resulting in identity of rat with human and/or mouse CXCR-4 (see Fig. 2). Constructs A through G were obtained by substitution of rat and human CXCR-4 at one or two of the conserved restriction sites *Bst*II, *Hinc*II, and *Apa*LI (see Fig. 2 and 3). Constructs H, I, and J, were obtained by ligation of PCR fragments amplified from rat and human CXCR-4 into R<sub>c</sub>/CMV between the *Hind*III and *Not*I sites. The rat CXCR-4 fragments were amplified with T7 (5'-AATACGA CTA CACTATAGG) as the 5' primer and either Nru-1 (5'-CGAAGATGATGT CAGGG), Mfe-1 (5'-TGGAACACCACCATCC), or Cla-1 (5'-TCGATGCTG ATCCCC) as the 3' primer. The human CXCR-4 fragments were amplified with either Nru-2 (5'-CGAACGTCAGTGAGGCAG), Mfe-2 (5'-ATTGCAGCACA TCATGGTTGG), or Cla-2 (5'-TTCCTTCATCCTCCTGG) as the 5' primer and SP6 (5'-ATTTAGGTACACTATAG) as the 3' primer. Constructs H to J were obtained by blunt-end ligation of PCR fragments T7/Nru-1 and Nru-2/SP6, T7/Mfe-1 and Mfe-2/SP6, and T7/Cla-1 and Cla-2/SP6, respectively, creating the restriction sites *Nru*I at the TM4/e3 junction (construct H), *Mfe*I at the e3/TM5 junction (construct I), and *Cla*I at the TM6/e4 junction (construct J). Construct K was derived by substitution of the *Apa*LI-*Not*I fragment from rat CXCR-4 into construct J. Construct L was obtained by blunt-end ligation of PCR fragments T7/Cla-1, amplified from construct A, and Cla-2/SP6, amplified from construct C. Construct M was obtained by blunt-end ligation of PCR fragments T7/Mfe-1, amplified from construct H, and Mfe-2/SP6, amplified from rat CXCR-4. Amino acid changes in the resulting protein, introduced by primers Mfe-1 and Mfe-2 (V194M and P198L), were shown to be detrimental to the coreceptor function and were corrected by site-directed mutagenesis with the oligonucleotide 5'-CA TGATGTGCTGAACTGGAACACAACCCACCAAGTCATT on a single-stranded DNA template. Construct N was obtained by blunt-end ligation of PCR fragments amplified from human CXCR-4 with primers T7 and Nru-4 (5'-CGAAGATGAAGTCGGGAA) and from construct I with primers Nru-3 (5'-GGAATGTCAGCCAGGGGG) and SP6. In this chimera, the first amino acid of e3 corresponds to the sequence of human CXCR-4 (N176) and not to rat CXCR-4 (D173).

**Mutations in human CXCR-4.** The Δ4 mutant was obtained by blunt-end ligation between the *Dra*I site (position 925 relative to the initial ATG codon) and a *Sma*I site of the pUC18 polylinker, creating a frameshift in the carboxy-terminal domain after F309, with the addition of 10 irrelevant amino acids. The N11Q and Δ4-41 mutants were obtained by site-directed mutagenesis on a single-stranded human CXCR-4 template with the oligonucleotides 5'-GGTGT ATTGATCAGAAGTGT (*Bcl*I site underlined) and 5'-GTAGATGGTGGGC CCTCCATGGT (*Bsp*120I site underlined), respectively. The Δ4-36 mutant was obtained by ligation of the partly complementary oligonucleotides 5'-CAT GGAGGGGAATAAGATCTTCC and 5'-CCGAGGAAGATCTTATTCCT into the Δ4-41 mutant digested with *Nco*I (position -2 relative to the initial ATG codon) and *Bsp*120I. The Δ2-9 mutant was obtained by digestion of the N11Q mutant with *Nco*I and *Bcl*I, filling in with T4 RNA polymerase, and blunt-end ligation.

**Functional assays for CXCR-4.** The U373MG-CD4 cells were transfected with wild-type or mutant CXCR-4 expression vectors in 24-well plates by calcium phosphate precipitation (overnight contact). Infections or cocultures were performed 24 h after transfection. The virus inoculum (10 to 30 ng of p24 antigen per well) was left in contact with cells for 36 to 40 h. Cocultures with Env<sup>+</sup> cells (1:1 ratio) were grown for 24 h. Cells were then fixed and stained for β-galactosidase activity with the X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) substrate, as described previously (13). Blue-stained cells or foci were scored under a magnification of ×20.

**Flow cytometry.** COS cells were transfected with expression vectors for CXCR-4 and green fluorescent protein (GFP) (EGFP-N1 vector; Clontech, Palo Alto, Calif.) by calcium phosphate precipitation. Cells were detached with phosphate-buffered saline containing 1 mM EDTA and pelleted 36 to 48 h after transfection. About 10<sup>6</sup> cells were incubated for 1 h at 4°C with monoclonal antibody 12G5 (kindly provided by J. Hoxie) at a concentration of 7.5 μg/ml in phosphate-buffered saline containing 2% bovine serum. Cells were washed three times and incubated for 1 h with phycoerythrin-conjugated rabbit anti-mouse immunoglobulins (Dako, Glostrup, Denmark). Stained cells were washed, fixed in 2% paraformaldehyde, and analyzed on an Epics Elite flow cytometer (Coultronics).

## RESULTS

**Strain specificity of rat CXCR-4.** The ability of wild-type or mutant chemokine receptors to behave as HIV coreceptors was tested by their expression in the CD4<sup>+</sup> human glioma cell line U373MG-CD4, which is naturally resistant to HIV-1 entry

and to fusion with Env<sup>+</sup> cells (19). This cell line was stably transfected with a Tat-inducible *lacZ* reporter gene (LTR<sub>lacZ</sub>), allowing us to detect complementation for HIV infection and for fusion with Tat<sup>+</sup> Env<sup>+</sup> cells by a simple in situ β-galactosidase assay, as described previously (19, 33).

Upon expression of human CXCR-4 or its rat homolog, U373MG-CD4 cells were susceptible to infection with HIV-1<sub>LAI</sub> and fusion with cells stably expressing the LAI Env proteins (Fig. 1). In contrast, only human CXCR-4 allowed infection of U373MG-CD4 cells with HIV-1<sub>NDK</sub> or HIV-2<sub>ROD</sub> and fusion with cells stably expressing chimeric LAI Env with the V3 loop of ROD or NDK (Fig. 1). In all subsequent experiments, parallel results were obtained in infection assays with NDK or ROD and in syncytium formation assays with HeLa cells stably expressing chimeric LAI Env with the V3 loop of NDK or ROD.

**Human-rat CXCR-4 chimeras.** Human and rat CXCR-4 have more than 90% amino acid identity. The differences are found essentially in the extracellular domains, in particular e1 and e3, while the TM and intracellular (i) domains are highly conserved or identical (Fig. 2). A first set of chimeras (constructs A to G) was constructed by using restriction sites conserved between the rat and human CXCR-4 cDNAs. The transfection of these chimeras into U373MG-CD4 cells allowed infection with HIV-1<sub>NL4.3</sub> (a recombinant HIV-1 strain with LAI env) (Table 1) and fusion with HeLa cells stably expressing LAI Env (Fig. 3b). Infection mediated by chimeras A and E was less efficient while infection mediated by chimeras D and F, both with the TM7 and i4 domains from rat CXCR-4, was more efficient than infection mediated by human CXCR-4. These differences were not strain specific and did not seem to be supported by differences in cell surface expression (Fig. 3c). This first series of constructs allowed us to rule out a role for genetic differences in the e1 to TM3 and i4 domains in the selectivity of rat CXCR-4 for LAI and in the ability of human CXCR-4 to mediate infection by NDK or ROD. Accordingly, chimeric rat CXCR-4 with the i2 to e4 domains of human CXCR-4 (construct F) had the same phenotype as human CXCR-4. Since the i2, i3, and TM5 domains of human and rat CXCR-4 are identical, the phenotypic differences between human and rat CXCR-4 can be supported only by genetic variation in the TM4, e3, TM6, and e4 domains.

The role of these domains was addressed with another series of chimeric receptors (constructs H to J) with replacements at the TM4/e3, e3/TM5, and TM6/e4 junctions (Fig. 3). Only construct H was able to mediate the infection of U373MG-CD4 cells by NDK or ROD (Table 1) and their fusion with HeLa cells expressing chimeric LAI Env (Fig. 3b). However, its efficiency was markedly reduced relative to that of infection mediated by human CXCR-4 or construct F. The comparison of results obtained with chimeras H and I indicates that the e3 domain of human CXCR-4 was necessary for NDK and ROD infection, while experiments with constructs K and L showed that differences in the e4 domain had no apparent role in mediating infection (Table 1). The substitution of the e3 domain from rat CXCR-4 into human CXCR-4 (construct N) abolished the coreceptor activity of CXCR-4 for NDK but not for LAI (Table 2). The reciprocal chimera (rat CXCR-4 with human e3 [construct M]) was able to mediate infection of U373MG-CD4 cells with LAI (Table 2), NDK (Table 2), and ROD (data not shown). It also allowed fusion with cells expressing LAI Env with NDK V3 (Fig. 3b). The efficiency of NDK infection was lower (about 50%) when U373MG-CD4 cells expressed chimera M than when they expressed human CXCR-4. However, chimera M was able to mediate NDK infection much more



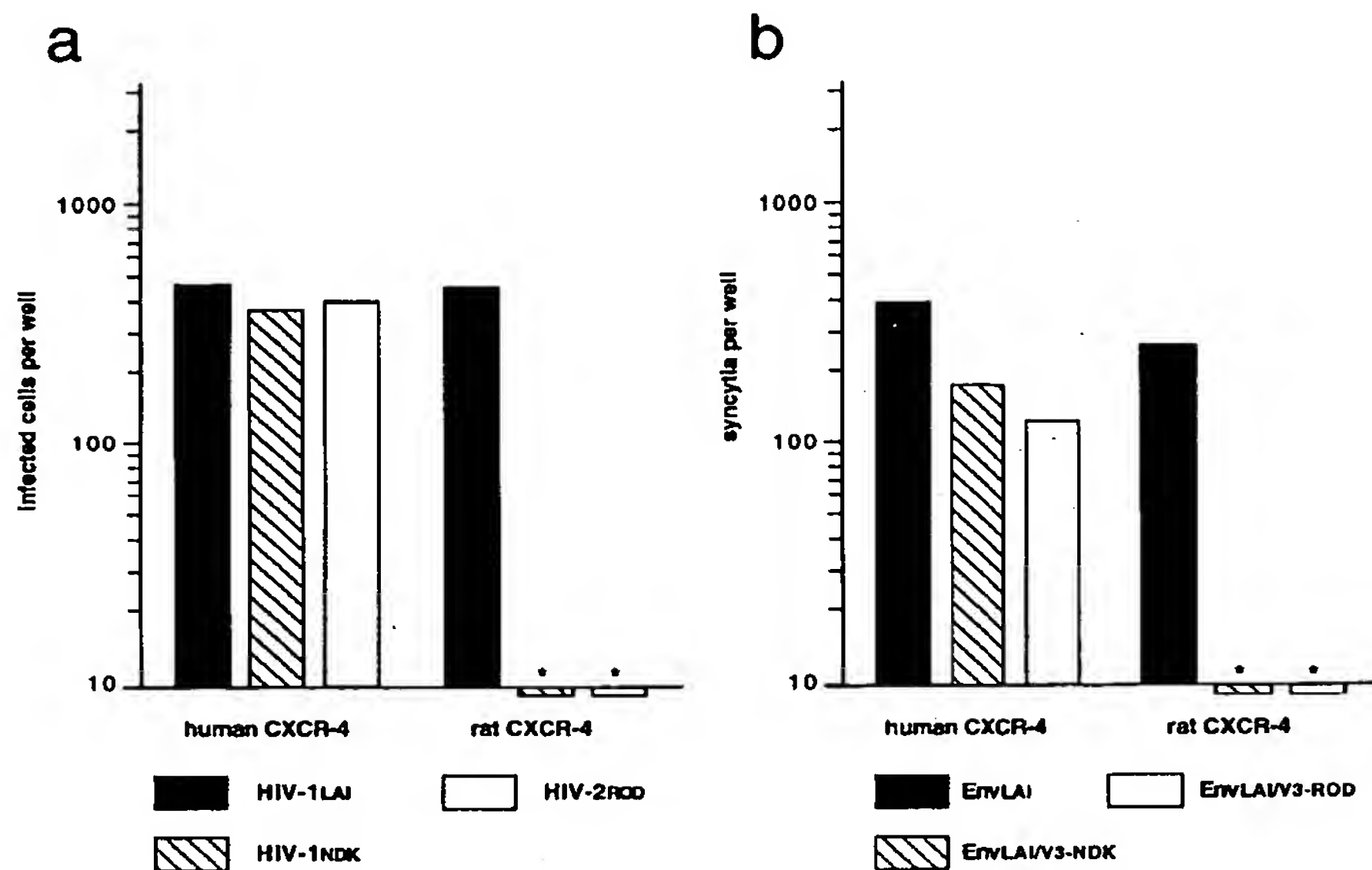


FIG. 1. Coreceptor activity of human and rat CXCR-4 for HIV-1 and HIV-2 strains. Target cells (U373MG-CD4 [LTRlacZ]) transfected with the indicated CXCR-4 expression vector were infected with HIV-1<sub>LAI</sub>, HIV-1<sub>NDK</sub>, or HIV-2<sub>ROD</sub> (a) or cocultured with HeLa cell lines stably expressing LAI Env or chimeric LAI Env with V3 from NDK or ROD (b). Infection with HIV or fusion with Env<sup>+</sup> cells results in transactivation of the LTRlacZ reporter gene and high  $\beta$ -galactosidase activity. Cells were stained with the X-Gal substrate 36 to 40 h after infection or 24 h after coculture. Bars represent the average numbers of blue-stained foci (log scale) in duplicate wells (24-well plate). Asterisks indicate that fewer than 10 foci were obtained.

efficiently than was chimera H. Also, its surface expression was apparently lower than that of chimera H or human CXCR-4 (Fig. 3c). The functional differences between chimeras H and M could be due to conformational effects

induced by chimerization. Overall, these experiments indicated that differences in the e3 domain were sufficient to account for the lack of coreceptor activity of rat CXCR-4 for NDK and ROD.

		<b>e1</b>	<b>TM1</b>	<b>i1</b>			
human	CXCR-4	MEGIS..IYTS	DNYTEEMGSGDYDSMK	PCFREENANFNKIFLPTIYSII	FLTGI	VGNGLVILVMGYQKKLR	70
rat	CXCR-4	-----S-	V-----N-----D-E--	R-----F-----			67
mouse	CXCR-4	--P--VS-	-----S-V-----N-----D-VH--	R-----F-----			72
						<u>BstE II</u>	
		<b>i1</b>	<b>TM2</b>	<b>e2</b>	<b>TM3</b>	<b>i2</b>	
human	CXCR-4	SMTDKYRLHL	SVADLLFVITLPFWAVDAVANWYFGN	FLCKAVHVIYTVNLYSSVLILAFISLD	RYLAIVHAT		142
rat	CXCR-4	-----M-D--	K-----I-----				139
mouse	CXCR-4	-----M-D--	K-----I-----				144
			***			<u>Hind II</u>	
		<b>i2</b>	<b>TM4</b>	<b>e3</b>	<b>TM5</b>		
human	CXCR-4	NSQRRPKLLAEKVVVVGVWIPALLLTIPDFIFANVSEA....	DDRYICDRFYPNDLWVVVFQFOHIMVGLI				209
rat	CXCR-4	-----A-----I--D-QG....-G-----L-DS-M-----					206
mouse	CXCR-4	-----A-----D-QGDISQG-----L-DS-M-----					216
		**		<u>Nru I</u>	<u>Mfe I</u>		
		<b>TM5</b>	<b>i3</b>	<b>TM6</b>	<b>e4</b>		
human	CXCR-4	LPGIVILSCYCIILISKLSHSGHQRKALKTTVILILAFFACWLPYYIGISIDSFILLEIIKQGC	EFENTVH				281
rat	CXCR-4	-----V-----V-----SV--					278
mouse	CXCR-4	-----V-----GV-----D--SI--					288
				<u>Cla I</u>	<u>ApaI I</u>		
		<b>TM7</b>	<b>i4</b>				
human	CXCR-4	KWISITEALAFFHCCLNPILYAF	LGA	FKTSAQH	ALTSVSRGSSLKILSKGKRGHSSVSTESESSSFHSS		352
rat	CXCR-4	-----S-----N-M-----					349
mouse	CXCR-4	-----S-----N-M-----					359

FIG. 2. Alignment of the amino acid sequences of human, rat, and mouse CXCR-4. The human and rat CXCR-4 sequences were determined directly from the expression vectors used in this study. Asterisks indicate discrepancies with the previously established sequence of rat CXCR-4 (42). The mouse CXCR-4 sequence is from the work of Nagasawa et al. (29). The predicted TM domains are underlined. The positions of the restriction sites used for construction of rat-human chimeras are indicated. Natural restriction sites are underlined.

TABLE 1. Infection of U373MG-CD4 (LTRlacZ) cells expressing WT or chimeric CXCR-4 by selected HIV strains

Construct used for transfection	No. of blue-stained cells/well <sup>a</sup> infected by:		
	HIV-1 <sub>NL4.3</sub>	HIV-1 <sub>NDK</sub>	HIV-2 <sub>ROD</sub>
None (mock)	1, 2	2, 3	2, 4
Human CXCR-4	450, 495	330, 400	410, 360
Rat CXCR-4	400, 460	1, 2	1, 0
Chimeras			
A	105, 90	75, 65	40, 50
B	280, 310	2, 1	2, 3
C	610, 650	2, 4	1, 0
D	1,250, 1,040	590, 530	410, 470
E	115, 105	0, 0	1, 0
F	810, 980	490, 550	320, 280
G	440, 530	2, 0	4, 1
H	340, 300	31, 36	30, 38
I	490, 570	2, 0	3, 2
J	430, 510	2, 0	2, 1
K	520, 550	2, 4	7, 4
L	650, 730	410, 480	520, 400

<sup>a</sup> Values for duplicate wells (24-well plates). Cells were stained with X-Gal for  $\beta$ -galactosidase activity 40 h after infection. All constructs were tested in the same experiment. Cell counts of >150 were obtained by extrapolation from randomly selected fields.

**Deletion mutants.** A series of deletions was engineered in human CXCR-4 cDNA by using natural restriction sites or site-directed mutagenesis. The deletion of the extracellular (e3 and e4) or intracellular (i2 and i3) loops abolished HIV coreceptor activity, but epitope-tagged forms of these mutants were not detected at the cell surface (data not shown), suggesting that these deletions were not compatible with correct folding of CXCR-4. The deletion of the carboxy-terminal intracytoplasmic domain ( $\Delta$ i4) was fully compatible with coreceptor activity for LAI, NDK, and ROD (Table 3). Higher numbers of infected cells, or syncytia with Env<sup>+</sup> cells, were observed when U373MG-CD4 cells were transfected with this mutant than with wild-type (WT) human CXCR-4.

The other mutations were introduced in the amino-terminal domain (e1) of human CXCR-4. The N11Q mutation, which suppressed one of the two potential N glycosylation sites in CXCR-4, increased its coreceptor activity for LAI, NDK, and ROD (Fig. 4; Table 3). The other potential N glycosylation site, at the TM4/e3 boundary (N176), is not conserved in rat CXCR-4. It seems, therefore, that N glycosylation is not necessary for HIV coreceptor activity. The deletion of the first nine amino acids of e1 (yielding the  $\Delta$ 2-9 mutant) had no apparent effect on infection or cell fusion mediated by LAI Env and slightly reduced the efficiency of NDK and ROD infection (Table 3). Surprisingly, CXCR-4 with an almost complete deletion of e1 ( $\Delta$ 4-36 mutant) was able to mediate LAI infection and the fusion of U373MG-CD4 cells with HeLa cells expressing LAI Env (Fig. 4; Table 3). This mutant was less efficient than WT CXCR-4, but flow cytometry experiments also showed reduced cell surface expression (see below). The coreceptor activity of the  $\Delta$ 4-36 mutant was extremely weak for NDK and null for ROD (Table 3), which is more evidence that these strains and LAI have different requirements for their interaction with CXCR-4. The  $\Delta$ 4-41 mutant mediated neither infection by LAI, NDK, and ROD nor fusion with Env<sup>+</sup> cells (Fig. 4; Table 3). However, an epitope-tagged form was not detected at the cell surface (data not shown), suggesting that

the deletion, which extends beyond the predicted e1/TM1 boundary, was not compatible with correct folding of CXCR-4.

**Mapping of the 12G5 epitope.** Monoclonal antibody 12G5 was selected for its ability to block cell fusion mediated by HIV-2 Env and was later found to detect a conformational epitope of human CXCR-4 (15). It was also recently shown to neutralize infection by HIV-1 and HIV-2 strains (26). Flow cytometry experiments were performed on simian COS cells cotransfected with expression vectors for human or rat CXCR-4 and for GFP. More than 80% of cells transfected with human CXCR-4 (GFP positive) were stained with 12G5, while about 2% of cells transfected with rat CXCR-4 were stained (Fig. 3c). Therefore, the 12G5 epitope is not conserved in rat CXCR-4. The rat-human CXCR-4 chimeras were tested for their reactivity with 12G5; it stained only cells expressing chimeras bearing the e3 domain of human CXCR-4, including construct M, in which e3 was the only domain from human CXCR-4. This indicated that the 12G5 epitope is located within the e3 domain of CXCR-4. However, it cannot be ruled out that the 12G5 epitope is formed by residues in e3 and in another domain conserved between human and rat CXCR-4. The percentage of stained cells and the mean fluorescence intensity were lower for cells transfected with chimera M than for cells expressing WT CXCR-4 (Fig. 3c and 5d). This could be due to reduced surface expression of the chimera or to less efficient presentation of the epitope in this context. Cells expressing the  $\Delta$ 4-36 mutant were stained by monoclonal antibody 12G5, although less efficiently than cells expressing WT CXCR-4 (Fig. 5f).

## DISCUSSION

The CXCR-4 chemokine receptor behaves as a coreceptor for HIV-1 and HIV-2 strains adapted to replicate in immortalized cell lines (15, 16, 33). It is also a coreceptor for dual-tropic HIV-1 strains (9, 12), for primary HIV-1 strains with an SI phenotype (37, 44), and for certain HIV-2 strains (38), all of which can also use CCR-5 as their coreceptor. The rat homolog of CXCR-4, differing from human CXCR-4 by a number of residues in the extracellular domains, is a coreceptor for the cell line-adapted strain LAI but not for the cell line-adapted strain NDK or for ROD (which also uses CCR-5). A series of chimeric molecules was constructed from human and rat CXCR-4 and functionally tested by its expression in a CD4<sup>+</sup> cell line (U373MG-CD4) naturally resistant to infection by HIV-1 and HIV-2 and to Env-mediated cell-cell fusion. All chimeras tested allowed infection by LAI, indicating that they were folded and expressed at the cell surface, which was confirmed by flow cytometry experiments. In the context of chimeras, the third extracellular domain (e3) of human CXCR-4 was both necessary and sufficient to allow observation of coreceptor activity for NDK and ROD. The simpler interpretation of these results is that e3 participates in the interaction of human CXCR-4 with NDK and ROD but not with LAI. Alternatively, these strains might interact differently with e3, with only LAI being able to tolerate the amino acid differences between human and rat CXCR-4. It seems less likely that e3 substitutions can have indirect effects and modulate the interaction of HIV with other domains of CXCR-4, although this possibility cannot be formally ruled out.

Monoclonal antibody 12G5 (15) stained cells expressing human, but not rat, CXCR-4. It reacted only with chimeras containing the e3 domain of human CXCR-4, indicating that e3 bears the 12G5 epitope or part of it. Antibody 12G5 can neutralize infection by HIV-1 and HIV-2 strains, but important variations depending upon the viral strains and cell types used

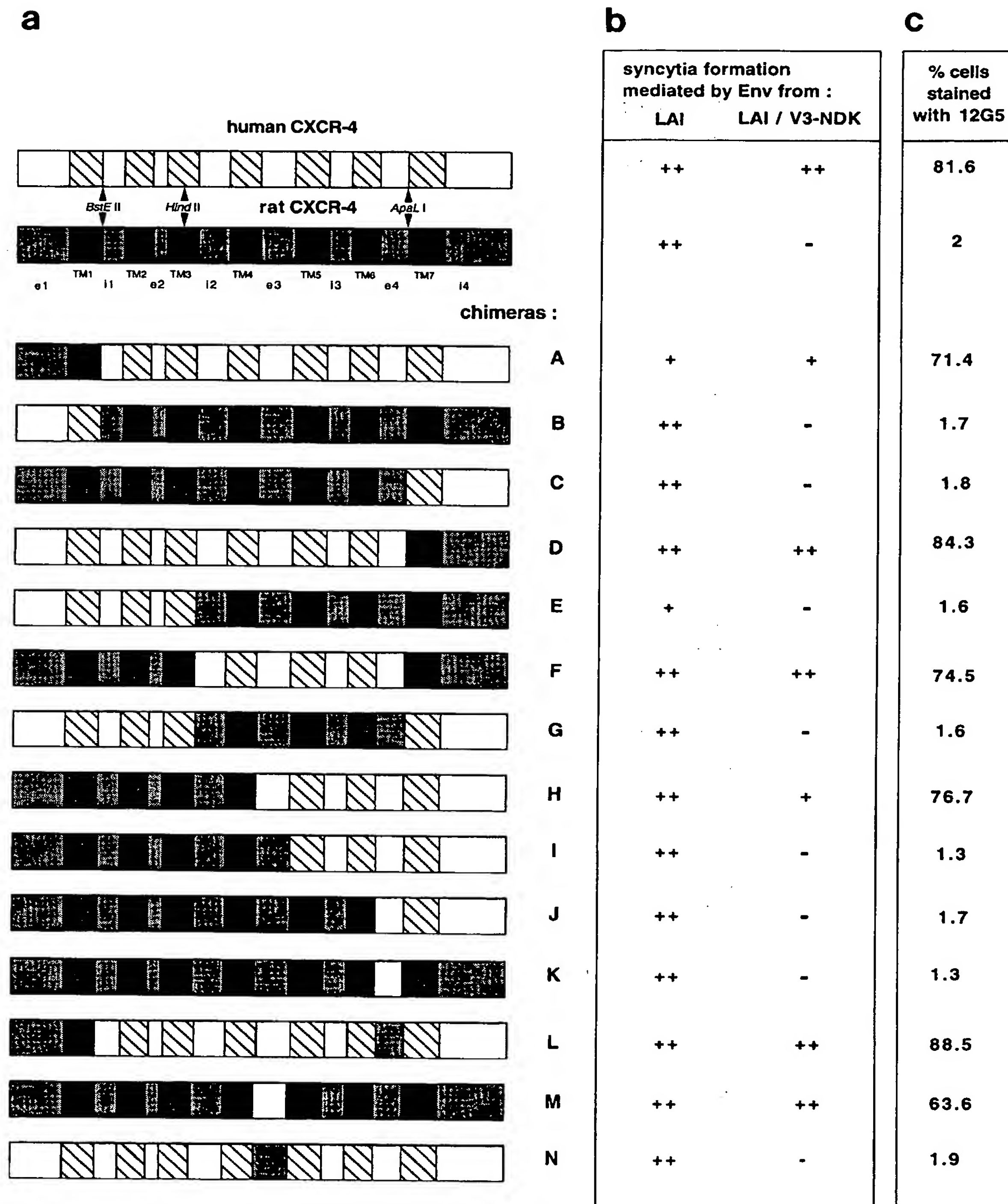


FIG. 3. (a) Diagrams of CXCR-4 constructs; (b) functionality of chimeric constructs as HIV-1 coreceptors; (c) construct reactivity with antibody 12G5. Chimeric constructs were transfected into U373MG-CD4 cells, and syncytium formation assays with HeLa cells expressing WT or chimeric LAI Env were performed as described in the legend to Fig. 1. Results are shown semiquantitatively, relative to the number of syncytia observed in parallel transfections with WT human CXCR-4. Symbols: ++, >60%; +, 20 to 60%; -, <10%. To assess 12G5 reactivity, COS cells were stained 36 h after cotransfection with WT or chimeric CXCR-4 and GFP vectors and analyzed by flow cytometry as described in the legend to Fig. 5. The percentage of 12G5-stained cells among GFP-positive cells was determined.



TABLE 2. Infection of U373MG-CD4 cells expressing WT or chimeric CXCR-4 by two HIV-1 strains

Construct used for transfection	No. of blue-stained cells/well <sup>a</sup> infected by:	
	HIV-1 <sub>LAI</sub>	HIV-1 <sub>NDK</sub>
Human CXCR-4	2,400, 2,400	2,000, 2,000
Rat CXCR-4	2,400, NA <sup>b</sup>	29, 28
Chimera M	1,300, 1,300	640, 600
Chimera N	1,300, 1,200	26, 20

<sup>a</sup> Values for duplicate wells. This experiment was performed as described in Table 1, footnote a. The larger numbers of infected cells relative to those in Tables 1 and 3 are due to a higher transfection efficiency.

<sup>b</sup> NA, not available.

seem to exist (26). Overall, it seems that cell line-adapted HIV-1 strains, such as LAI, are relatively resistant to neutralization by 12G5 compared to dual-tropic HIV-1 strains or to HIV-2 (26). If this result is confirmed, it could also suggest that the e3 domain is not required for the interaction of LAI with CXCR-4. However, the antiviral activity of 12G5 could have other mechanisms. The binding of 12G5 might exert steric hindrance effects on domains other than e3, or it might even induce CXCR-4 internalization. In that case, strains such as LAI might be less affected than others by the reduced expression of their coreceptor at the cell surface. At this time, it is difficult to sort out these possible mechanisms supporting the antiviral activity of antibody 12G5.

The mutation of a potential N glycosylation site in the amino-terminal domain (e1) increased the coreceptor activity of CXCR-4 for all strains tested. It can be envisioned that the absence of N-linked sugars facilitates coreceptor access. N glycosylation was also found to be dispensable for the coreceptor activity of CCR-5 (34). The truncation of the carboxy-terminal intracytoplasmic domain (i4) increased the coreceptor activity of CXCR-4. The i4 domain contains a number of

TABLE 3. Infection of U373MG-CD4 cells expressing WT or mutant human CXCR-4 by selected HIV strains

Construct used for transfection	No. of blue-stained cells/well <sup>a</sup> infected by:		
	HIV-1 <sub>NL4.3</sub>	HIV-1 <sub>NDK</sub>	HIV-2 <sub>ROD</sub>
None (mock)	2, 2	1, 3	3, 5
WT CXCR-4	480, 530	460, 480	320, 270
CXCR-4 mutants			
Δi4	750, 660	820, 950	500, 620
N11Q	1,050, 1,250	850, 900	750, 810
Δ2-9	480, 560	370, 440	250, 220
Δ4-36	120, 150	5, 20	2, 3
Δ4-41	5, 2	7, 3	0, 0

<sup>a</sup> Values for duplicate wells. This experiment was performed as described in Table 1, footnote a.

serine and threonine residues representing targets for phosphorylation events modulating the activity of G-protein-coupled receptors (28). The coupling of CXCR-4 via its i4 domain is therefore dispensable for HIV coreceptor activity. The increased activity of the Δi4 mutant was probably an indirect effect of the i4 deletion on, for example, the lateral mobility or the turnover of CXCR-4 at the cell surface. The intracellular loops i2 and i3 are also involved in coupling to the signal transduction machinery, but their deletion was not compatible with the surface expression of CXCR-4, due probably to severe conformational effects. A more detailed mutagenesis study of these domains will be required to address their possible role in HIV coreceptor activity.

The deletion of the first nine amino acids of e1 had no apparent effect on the coreceptor activity of CXCR-4 for LAI and slightly reduced its activity for NDK and ROD. The almost complete deletion of e1 (yielding the Δ4-36 mutant) was compatible with the cell surface expression and coreceptor activity

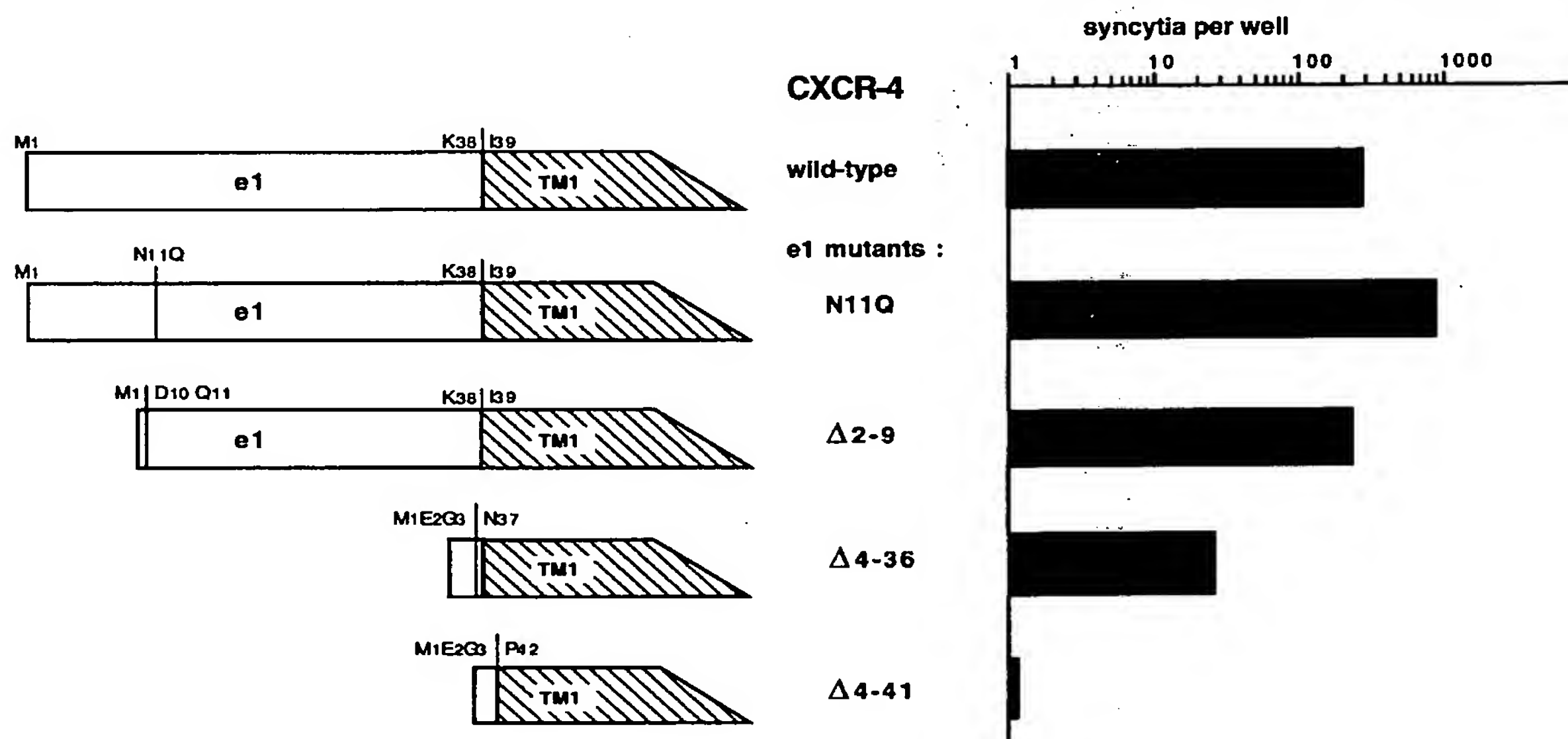


FIG. 4. Coreceptor activity of CXCR-4 with mutations or deletions in the amino-terminal e1 domain. Transfection of U373MG-CD4 cells and syncytium formation assays with HeLa cells stably expressing LAI Env were performed as described in the legend to Fig. 1.

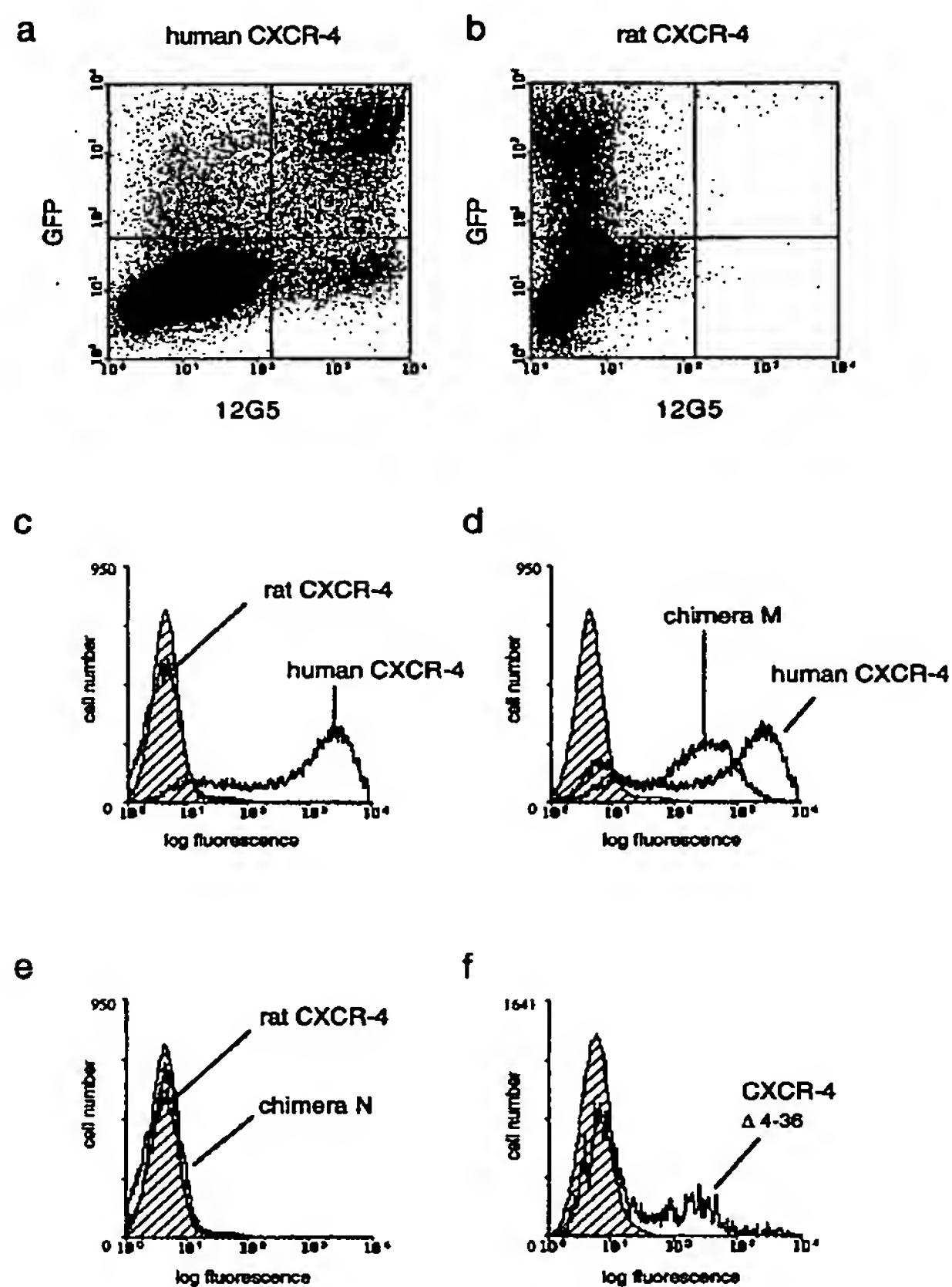


FIG. 5. Flow cytometry analysis of CXCR-4 expression. COS cells were transfected with a GFP expression vector and with WT or mutant CXCR-4, as indicated (1:6 ratio), and stained either with the anti-CXCR-4 monoclonal antibody 12G5 and a phycoerythrin-conjugated secondary antibody or with the secondary antibody only (shaded curves in panels c to f). (a and b) Two-color analysis of 12G5 (x axis, red fluorescence) and GFP (y axis, green fluorescence) expression (arbitrary units); (c to f) distribution of 12G5 expression among GFP-positive cells ( $>10^2$  arbitrary units) (x axis, red fluorescence; y axis, cell numbers).

of CXCR-4 for LAI, although both were reduced relative to those of WT CXCR-4. Notably, structure-function studies on the interleukin-8 receptor (CXCR-1) have suggested the existence of a disulfide bridge between cysteines of e1 and e4 (20). These cysteines are conserved in other chemokine receptors, including CXCR-4. Our results with the  $\Delta 4-36$  mutant indicate that the formation of a disulfide bridge between e1 and e4 is not absolutely necessary for the processing of CXCR-4 to the cell surface. The  $\Delta 4-36$  mutant had very low coreceptor activity for NDK and did not mediate infection with ROD. The e1 domain might therefore be dispensable for the interaction of CXCR-4 with gp120 of LAI but required for the other strains. It can also be envisioned that LAI is affected differently by the reduced expression of the  $\Delta 4-36$  mutant at the cell surface than is NDK or ROD. Similar studies with CCR-5 have shown that the e1 and e2 domains were necessary for coreceptor activity for primary and dual-tropic HIV-1 strains, while the e3 and e4 domains had no apparent role (4, 34). Therefore, differences in the interaction of dual-tropic HIV strains with CXCR-4 and CCR-5, as well as in the interaction of HIV strains with CXCR-4, seem to exist.

The HIV surface envelope protein gp120 seems to interact directly with the chemokine receptors CXCR-4 and CCR-5, this interaction being markedly reinforced by the presence of CD4 (22, 40, 43). The third hypervariable domain (V3) of gp120 was found to be necessary for the interaction of gp120 with CCR-5, although other domains of gp120 were apparently also involved in this interaction (40, 43). The role of V3 in the interaction of gp120 with CXCR-4 has not been directly observed but has been strongly suggested by different experiments. In cell line-adapted HIV-1, V3 substitutions resulted in a coreceptor switch from CXCR-4 to CCR-5 (9). We have observed that the V3 domain of LAI gp120 was required for utilization of the rat coreceptor CXCR-4 and conferred this property to chimeric NDK gp120 (33). A direct role for V3 in gp120-CXCR-4 interaction would be consistent with the greater sensitivity of cell line-adapted HIV-1 strains to neutralization by anti-V3 antibodies (reviewed in reference 35) or drugs targeting V3 (24, 31) compared with that of primary HIV-1 strains using the CCR-5 coreceptor. The V3 mutations associated with the switch of HIV-1 strains from the NSI to the SI phenotype usually increase the net positive charge of V3 (8, 17, 21). The e3 domain of CXCR-4 is negatively charged (net charge,  $-3$ ), while the homologous domain of CCR-5 is positively charged ( $+2$ ). It is tempting to speculate that electrostatic interactions take place between the V3 loop of SI strains and the e3 domain of CXCR-4. A detailed mutagenesis study of the charged residues in e3 could make it possible to address this hypothesis.

Even if a role for V3 in the interaction of gp120 with coreceptors is probable, the very slight similarity between HIV-1 and HIV-2 in this domain suggests a role for a more conserved domain(s) of gp120. It seems unlikely to us that gp120 can accommodate totally distinct sites, conserved across strains and primate immunodeficiency viruses, which mediate the interaction with CCR-5, CXCR-4, and possibly other chemokine receptors. A simpler view might be that the same domain, or conformation, of gp120 has the ability to interact with different chemokine receptors, possibly via their amino-terminal domains. This interaction might be sufficient for an HIV strain to use CCR-5, but not CXCR-4, as its coreceptor, thus explaining the preferential use of CCR-5 by primary NSI strains. Mutations in other domains of gp120 (in particular, V3) might be required for a functional interaction with CXCR-4, possibly involving the e3 domain. The HIV strains bearing these mutations (primary SI or dual tropic) could therefore use either CCR-5 or CXCR-4. In cell line-adapted strains, mutations in gp120 resulting in a tighter interaction with CXCR-4 might be detrimental to interaction with CCR-5, thus preventing the use of this coreceptor. This model is consistent with the observation that V3 mutations or substitutions conferring the NSI phenotype to a cell line-adapted HIV-1 strain were associated to a coreceptor switch from CXCR-4 to CCR-5 (9).

If a conserved domain of gp120 allows interaction with chemokine receptors as different as CCR-5 and CXCR-4, it can be wondered why other chemokine receptors do not behave as HIV coreceptors (CCR-1, CCR-4, or CCR-2a) or can be used only by certain viral strains (CCR-3 and CCR-2b). The ability to interact with gp120 might be necessary but not sufficient for HIV coreceptor activity, which may require other properties, such as the ability to colocalize with CD4 either spontaneously or in the presence of gp120.

#### ACKNOWLEDGMENTS

A. BreLOT and N. Heveker made equal contributions to this work. We thank our colleagues J. Hoxie (University of Pennsylvania, Philadelphia) and R. Duman (Yale University, New Haven, Conn.) for



generous gifts of reagents and I. Bouchaert and F. Letourneur (ICGM) for help with flow cytometry and sequencing.

This work was supported by the Agence Nationale de Recherches sur le SIDA and by Ensemble contre le SIDA.

## REFERENCES

- Adachi, A., H. E. Gendelman, S. Koenig, T. Folks, R. Willey, A. Rabson, and M. A. Martin. 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J. Virol.* 59:284-291.
- Alkhatib, G., C. Combadiere, C. C. Broder, Y. Feng, P. E. Kennedy, P. M. Murphy, and E. A. Berger. 1996. CC CKR5: a RANTES, MIP-1 $\alpha$ , MIP-1 $\beta$  receptor as a fusion cofactor for macrophage-tropic HIV-1. *Science* 272:1955-1958.
- Ashorn, P. A., E. A. Berger, and B. Moss. 1990. Human immunodeficiency virus envelope glycoprotein/CD4-mediated fusion of nonprimate cells with human cells. *J. Virol.* 64:2149-2156.
- Atchison, R. E., J. Gosling, F. S. Monteclaro, C. Franci, L. Digilio, I. F. Charo, and M. A. Goldsmith. 1996. Multiple elements of CCR5 and HIV-1 entry: dissociation from response to chemokines. *Science* 274:1924-1926.
- Bleul, C. C., M. Farzan, H. Choe, C. Parolin, I. Clark-Lewis, J. Sodroski, and T. A. Springer. 1996. The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry. *Nature* 382:829-833.
- Broder, C. C., D. S. Dimitrov, R. Blumenthal, and E. A. Berger. 1993. The block to HIV-1 envelope glycoprotein-mediated membrane fusion in animal cells expressing human CD4 can be overcome by a human cell component(s). *Virology* 193:483-491.
- Brun-Vézinet, F., M. A. Rey, C. Katlama, P. M. Girard, D. Roulot, P. Yéni, L. Lenoble, F. Clavel, M. Alizon, S. Gadelle, J. J. Madjar, and M. Harzic. 1987. Lymphadenopathy-associated virus type 2 in AIDS and AIDS-related complex. *Lancet* i:128-132.
- Callahan, L. 1994. HIV-1 virion-cell interactions: an electrostatic model of pathogenicity and syncytium formation. *AIDS Res. Hum. Retroviruses* 10:231-233.
- Choe, H., M. Farzan, Y. Sun, N. Sullivan, B. Rollins, P. D. Ponath, L. Wu, C. R. Mackay, G. LaRosa, W. Newman, N. Gerard, C. Gerard, and J. Sodroski. 1996. The  $\beta$ -chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell* 85:1135-1148.
- Clapham, P. R., D. Blanc, and R. A. Weiss. 1991. Specific cell surface requirements for the infection of CD4-positive cells by human immunodeficiency virus types 1 and 2 and by simian immunodeficiency virus. *Virology* 181:703-715.
- Deng, H., R. Liu, W. Elmeier, S. Choe, D. Unutmaz, M. Burkhart, P. Di Marzio, S. Marmon, R. E. Sutton, C. M. Hill, C. B. Davis, S. C. Peiper, T. J. Schall, D. R. Littman, and N. R. Landau. 1996. Identification of a major co-receptor for primary isolates of HIV-1. *Nature* 381:661-666.
- Doranz, B. J., J. Rucker, Y. Yi, R. J. Smyth, M. Samson, R. C. Peiper, M. Parmentier, R. G. Collman, and R. W. Doms. 1996. A dual-tropic primary HIV-1 isolate that uses fusin and the  $\beta$ -chemokine receptors CKR-5, CKR-3, and CKR-2B as fusion cofactors. *Cell* 85:1149-1158.
- Dragic, T., P. Charneau, F. Clavel, and M. Alizon. 1992. Complementation of murine cells for human immunodeficiency virus envelope/CD4-mediated fusion in human/murine heterokaryons. *J. Virol.* 66:4794-4802.
- Dragic, T., V. Litwin, G. Allaway, S. R. Martin, Y. Huang, K. A. Nagashima, C. Cayanan, P. J. Maddon, R. A. Koup, J. P. Moore, and W. A. Paxton. 1996. HIV-1 entry into CD4<sup>+</sup> cells is mediated by the chemokine receptor CC-CKR5. *Nature* 381:667-673.
- Endres, M. J., P. R. Clapham, M. Marsh, M. Ahuja, J. Davis Turner, A. McKnight, J. F. Thomas, B. Stoeckenau-Haggarty, S. Choe, P. J. Vance, T. N. C. Wells, C. A. Power, S. S. Sutterwala, R. W. Doms, N. R. Landau, and J. A. Hoxie. 1996. CD4-independent infection by HIV-2 is mediated by fusin/CXCR4. *Cell* 87:745-756.
- Feng, Y., C. C. Broder, P. E. Kennedy, and E. A. Berger. 1996. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane G-protein coupled receptor. *Science* 272:872-877.
- Fouchier, R. A. M., M. Groenink, N. A. Kootstra, M. Tersmette, H. G. Huisman, F. Miedema, and H. Schuitemaker. 1992. Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. *J. Virol.* 66:3183-3187.
- Freed, E. O., and M. A. Martin. 1995. The role of human immunodeficiency virus type 1 envelope glycoproteins in virus infection. *J. Biol. Chem.* 270:23883-23886.
- Harrington, R. D., and A. P. Geballe. 1993. Cofactor requirement for human immunodeficiency virus type 1 entry into a CD4-expressing human cell line. *J. Virol.* 67:5939-5947.
- Hébert, C. A., A. Chuntharapai, M. Smith, T. Colby, J. Kim, and R. Horuk. 1993. Partial functional mapping of the human interleukin-8 type A receptor. *J. Biol. Chem.* 268:18549-18553.
- Kuiken, C. L., J.-J. de Jong, E. Baan, W. Keulen, M. Tersmette, and J. Goudsmit. 1992. Evolution of the V3 envelope domain in proviral sequences and isolates of human immunodeficiency virus type 1 during transition of the viral biological phenotype. *J. Virol.* 66:4622-4627.
- Lapham, C. K., J. Ouyang, B. Chandrasekhar, N. Y. Nguyen, D. S. Dimitrov, and H. Golding. 1996. Evidence for cell-surface association between fusin and the CD4-gp120 complex in human cell lines. *Science* 274:602-605.
- Levy, J. A. 1993. Pathogenesis of human immunodeficiency virus infection. *Microbiol. Rev.* 57:183-289.
- Lynch, G., L. Low, S. Li, A. Sloane, S. Adams, C. Parish, B. Kemp, and A. L. Cunningham. 1994. Sulfated polyanions prevent HIV infection of lymphocytes but do not inhibit monocyte infection. *J. Leukocyte Biol.* 56:266-272.
- Maddon, P. J., A. G. Dalgleish, J. S. McDougal, P. R. Clapham, R. A. Weiss, and R. Axel. 1986. The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell* 47:333-348.
- McKnight, A., D. Wilkinson, G. Simmons, S. Talbot, L. Picard, M. Ahuja, M. Marsh, J. A. Hoxie, and P. R. Clapham. 1997. Inhibition of human immunodeficiency virus fusion by a monoclonal antibody to a coreceptor (CXCR4) is both cell type and virus strain dependent. *J. Virol.* 71:1692-1696.
- Moore, J. P., B. A. Jameson, R. A. Weiss, and Q. J. Sattentau. 1993. The HIV-cell fusion reaction, p. 233-289. *In* J. Bentz (ed.), *Viral fusion mechanisms*. CRC Press, Inc., Boca Raton, Fla.
- Murphy, P. M. 1994. The molecular biology of leucocyte chemoattractant receptors. *Annu. Rev. Immunol.* 12:593-633.
- Nagasawa, T., T. Nakajima, K. Tachibana, H. Iizasa, C. C. Bleul, O. Yoshie, K. Matsushima, N. Yoshida, T. A. Springer, and T. Kishimoto. 1996. Molecular cloning and characterization of a murine pre-B-cell growth-stimulating factor/stromal cell-derived factor 1 receptor, a murine homolog of the human immunodeficiency virus 1 entry coreceptor fusin. *Proc. Natl. Acad. Sci. USA* 93:14726-14729.
- Oberlin, E., A. Amara, F. Bachelier, C. Bessia, J.-L. Virelizier, F. Arenzana-Seisdenos, O. Schwartz, J.-M. Heard, D. F. Legler, M. Loetscher, M. Baggiolini, and B. Moser. 1996. The CXCR chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. *Nature* 382:823-835.
- O'Brien, W. A., M. Sumner-Smith, S.-H. Mao, S. Sadeghi, J.-Q. Zhao, and I. S. Y. Chen. 1996. Anti-human immunodeficiency virus type 1 activity of an oligocationic compound mediated via gp120 V3 interactions. *J. Virol.* 70:2825-2831.
- Peden, K., M. Emerman, and L. Montagnier. 1991. Changes in growth properties on passage in tissue culture of viruses derived from infectious molecular clones of HIV-1<sub>LAI</sub>, HIV-1<sub>MAL</sub>, and HIV-1<sub>ELI</sub>. *Virology* 185:661-672.
- Pleskoff, O., N. Sol, B. Labrosse, and M. Alizon. 1997. Human immunodeficiency virus strains differ in their ability to infect CD4<sup>+</sup> cells expressing the rat homolog of CXCR-4 (fusin). *J. Virol.* 71:3259-3262.
- Rucker, J., M. Samson, B. J. Doranz, F. Libert, J. F. Berson, Y. Yi, R. J. Smyth, R. G. Collman, C. C. Broder, G. Vassart, R. W. Doms, and M. Parmentier. 1996. Regions in  $\beta$ -chemokine receptors CCR5 and CCR2b that determine HIV-1 cofactor activity. *Cell* 87:437-446.
- Sattentau, Q. J. 1996. Neutralization of HIV-1 by antibody. *Curr. Opin. Immunol.* 8:540-545.
- Schwartz, O., M. Alizon, J. M. Heard, and O. Danos. 1994. Impairment of T cell receptor-dependent stimulation in CD4<sup>+</sup> lymphocytes after contact with membrane-bound HIV-1 envelope glycoprotein. *Virology* 198:360-365.
- Simmons, G., D. Wilkinson, J. D. Reeves, M. T. Dittmar, S. Beddows, J. Weber, G. Carnegie, U. Desselberger, P. W. Gray, R. A. Weiss, and P. R. Clapham. 1996. Primary, syncytium-inducing human immunodeficiency virus type 1 isolates are dual-tropic and most can use either Lestr or CCR5 as coreceptors for virus entry. *J. Virol.* 70:8355-8360.
- Sol, N., and M. Alizon. Unpublished data.
- Spire, B., J. Sire, V. Zachar, F. Rey, F. Barré-Sinoussi, F. Galibert, A. Hampe, and J.-C. Chermann. 1989. Nucleotide sequence of HIV1-NDK, a highly cytopathic strain of the human immunodeficiency virus, HIV1. *Gene* 81:275-284.
- Trkola, A., T. Dragic, J. Arthos, J. M. Binley, W. C. Olson, G. Allaway, S. R. Martin, C. Cheng-Mayer, J. Robinson, P. J. Maddon, and J. P. Moore. 1996. CD4-dependent, antibody-sensitive interactions between HIV-1 and its co-receptor CCR5. *Nature* 384:184-187.
- Weiss, R. A. 1993. Cellular receptors and viral glycoproteins involved in retrovirus entry, p. 1-108. *In* J. A. Levy (ed.), *The Retroviridae*, vol. 2. Plenum Press, New York, N.Y.
- Wong, M.-L., W. W. Xin, and R. S. Duman. 1996. Rat LCR1: cloning and cellular distribution of a putative chemokine receptor in the brain. *Mol. Psychiatry* 1:133-140.
- Wu, L., N. P. Gerard, R. Wyatt, H. Choe, C. Parolin, N. Ruffing, A. Borsetti, A. A. Cardoso, E. Desjardin, W. Newman, C. Gerard, and J. Sodroski. 1996. CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. *Nature* 384:179-183.
- Zhang, L., Y. Huang, T. He, Y. Cao, and D. D. Ho. 1996. HIV-1 subtype and second-receptor use. *Nature* 383:768.



## In Vitro Mutagenesis Identifies a Region within the Envelope Gene of the Human Immunodeficiency Virus That Is Critical for Infectivity

RONALD L. WILLEY,<sup>1\*</sup> DOUGLAS H. SMITH,<sup>2</sup> LAURENCE A. LASKY,<sup>2</sup> THEODORE S. THEODORE,<sup>1</sup>  
PATRICIA L. EARL,<sup>3</sup> BERNARD MOSS,<sup>3</sup> DANIEL J. CAPON,<sup>2</sup> AND MALCOLM A. MARTIN<sup>1</sup>

*Laboratories of Molecular Microbiology<sup>1</sup> and Viral Diseases,<sup>3</sup> National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892, and Department of Molecular Biology, Genentech, Inc., South San Francisco, California 94080<sup>2</sup>*

Received 11 August 1987/Accepted 25 September 1987

Site-specific mutagenesis was used to introduce amino acid substitutions at the asparagine codons of four conserved potential N-linked glycosylation sites within the gp120 envelope protein of human immunodeficiency virus (HIV). One of these alterations resulted in the production of noninfectious virus particles. The amino acid substitution did not interfere with the synthesis, processing, and stability of the *env* gene polypeptides gp120 and gp41 or the binding of gp120 to its cellular receptor, the CD4 (T4) molecule. Vaccinia virus recombinants containing wild-type or mutant HIV *env* genes readily induced syncytia in CD4<sup>+</sup> HeLa cells. These results suggest that alterations involving the second conserved domain of the HIV gp120 may interfere with an essential early step in the virus replication cycle other than binding to the CD4 receptor. In long-term cocultures of a T4<sup>+</sup> lymphocyte cell line and colon carcinoma cells producing the mutant virus, revertant infectious virions were detected. Molecular characterization of two revertant proviral clones revealed the presence of the original mutation as well as a compensatory amino acid change in another region of HIV gp120.

The acquired immunodeficiency syndrome (AIDS) is caused by a novel human retrovirus (4, 11, 21) known as the human immunodeficiency virus (HIV). In tissue culture systems, HIV infects and kills human CD4<sup>+</sup> (T4<sup>+</sup>) lymphocytes and lymphocyte lines (17). In infected individuals, the number of circulating T4<sup>+</sup> lymphocytes may ultimately decline; this reduction is usually correlated with immunologic dysfunction (19).

Previous studies have shown that HIV and the T4A monoclonal antibody can reciprocally compete in binding to T4<sup>+</sup> lymphocytes (25). Analysis of this reaction subsequently revealed that it involved the direct interaction of the HIV gp120 with a specific portion of the CD4 (T4) molecule (24, 26). The functional significance of CD4 as a receptor for HIV has recently been demonstrated by introducing CD4 cDNA into nonlymphocyte human cell lines (23). The subsequent expression of the CD4 antigen on the surface of such cells renders them susceptible to HIV infection.

Viral envelopes have also been shown to participate in early events in the replicative cycle that follow the binding of particles to their receptor(s) (8). For example, a fusion step is required for passage through the two lipid-bilayer membranes (in the virus particle and the cell surface) which separate the viral genome from the cytoplasm of the infected cell. Sendai virus encodes an inactive precursor protein (F0) which must be cleaved to the mature F1 and F2 products that mediate, via a fusion mechanism, the direct passage of particles into cells (15, 33). Virus fusion with the host membrane also appears to be the mode of entry for HIV (36).

Analyses of the *env* genes present in different HIV isolates have indicated that they contain several highly variable and conserved domains (3, 35, 41). In some African isolates, greater than 50% variability was noted at the deduced amino

acid level involving stretches of more than 100 residues of gp120 (3, 41). Despite this degree of *env* heterogeneity, all HIV isolates tested have retained their tropism for T4<sup>+</sup> lymphocytes and their ability to induce cytopathic effects during productive infection in tissue culture. This suggests that one or more of the conserved domains within gp120 are involved in the adsorption of HIV to the CD4 molecule and subsequent steps in viral replication.

In this work, we have focused on the second conserved HIV *env* domain (41) and used site-specific mutagenesis techniques to target potential N-linked (Asn-X-Ser/Thr) glycosylation sites, where X is any amino acid. One of four HIV mutants examined was rendered noninfectious by this procedure. In vitro assays indicated that the gp120 produced by this mutant bound to the CD4 molecule. This result implies that portions of the HIV envelope not involved in binding to CD4 may be necessary for the initiation of productive viral infection. Interaction with other cellular surface proteins or processing of the bound virions may be required subsequent to their interaction with the CD4 receptor.

### MATERIALS AND METHODS

**Oligonucleotide synthesis and purification.** Oligodeoxyribonucleotides were synthesized on an Applied Biosystems model 380B DNA synthesizer by using phosphoramidite chemistry. Oligonucleotides were purified on 20% polyacrylamide gels and then with a Sep-Pak C18 column (Waters Associates) as described previously (31).

**Oligonucleotide-directed mutagenesis.** Mutants were derived from an infectious plasmid clone of HIV, pNL4-3 (1). A 2.7-kilobase (kb) *EcoRI*-*Bam*HI restriction fragment containing the amino-terminal 2,267 nucleotides of the envelope (*env*) gene was cloned into M13 vectors mp18 and mp19, and specific nucleotide changes were introduced by oligonucleotide-directed mutagenesis (42). The mutagenic oligonucle-

\* Corresponding author.

otides (designation) were 5'-ACCATGTACAcAaGTCAGC ACAG-3' (6992), 5'-ACTGCTGTTAcAaGGCAGTCTAG-3' (7055), 5'-TAGATCTGCCAAaTTCACAGACA-3' (7099), 5'-AGTACAGCTGcAaCAATCTGTAG-3' (7136), 5'-CAACT GCTGTTA(aat)GGCAGTCTAGCA-3' (7055D), 5'-CAACT GCTGgacAATGGCAGTC-3' (7052), and 5'-CTGTAAAT GaCAGTCTAGCA-3' (7058). The lowercase letters identify nucleotide changes that were introduced. In 7055D, letters in parentheses indicate nucleotides that were deleted. The different mutant 2.7-kb *EcoRI-BamHI* fragments were each isolated from the replicative form of M13 DNA and cloned back into *EcoRI-BamHI*-restricted pNL4-3 DNA to generate mutant HIV plasmid clones p6992, p7055, p7099, p7136, p7055D, p7052, and p7058.

**Cell cultures and DNA transfection.** A3.01, a human T4<sup>+</sup> lymphocytic leukemia cell line (10), was maintained in RPMI 1640 medium (GIBCO Laboratories) supplemented with 10% heat-inactivated fetal calf serum (FCS). Cells (approximately  $4 \times 10^6$ /ml) were transfected by electroporating (29) 20  $\mu$ g of uncleaved plasmid DNA. The transfected A3.01 cells were maintained in 25-cm<sup>2</sup> tissue culture flasks (Costar) containing 10 ml of RPMI-FCS. An adherent colon carcinoma cell line, SW480 (ATCC CCL228), was maintained in Dulbecco modified Eagle medium supplemented with 10% FCS. Cells were trypsinized, diluted, and transferred to 25-cm<sup>2</sup> flasks 14 to 16 h prior to transfection. Individual 60 to 80% confluent cultures were transfected with 10  $\mu$ g of uncleaved plasmid DNA by the calcium phosphate precipitation method (40). Cocultures were established by scraping transfected SW480 cells from individual 25-cm<sup>2</sup> flasks 24 to 72 h following transfection and adding them to  $2 \times 10^6$  A3.01 cells in 25-cm<sup>2</sup> flasks containing 10 ml of RPMI-FCS.

**Infections.** The infectivity of progeny virions generated from transfection experiments was assayed in the A3.01 lymphocyte cell line. Supernatants were filtered (0.22  $\mu$ m pore size; Millex) and 300  $\mu$ l was added to 15-ml conical tubes (Falcon) containing  $2 \times 10^6$  A3.01 cells, 700  $\mu$ l of RPMI-FCS, and 2  $\mu$ g of Polybrene (Sigma Chemical Co.). After 4 h at 37°C, the cells were washed twice with phosphate-buffered saline, suspended in 10 ml of RPMI-FCS, and maintained in 25-cm<sup>2</sup> culture flasks.

**RT assay.** Virion-associated reverse transcriptase (RT) activity was measured as described previously (13) with modifications as follows: 10  $\mu$ l of culture supernatant was mixed with 50  $\mu$ l of an RT reaction mixture which contained a template primer of (A)<sub>n</sub> (5  $\mu$ g/ml) and (dT)<sub>12-18</sub> (1.57  $\mu$ g/ml) in 50 mM Tris, pH 7.8, 7.5 mM KCl, 2 mM dithiothreitol, 5 mM MgCl<sub>2</sub>, 0.05% Nonidet P-40, and 0.5  $\mu$ Ci of [<sup>32</sup>P]dTTP (400 Ci/mmol). Following a 90-min incubation at 37°C, 10  $\mu$ l of the reaction mixture was spotted onto DEAE ion-exchange paper (Whatman) and washed four to six times in  $2 \times$  SSC (1  $\times$  SSC is 0.15 M NaCl plus 0.015 M sodium chloride) to remove unincorporated [<sup>32</sup>P]dTTP. Spots were visualized by autoradiography or counted in a scintillation counter.

The association of RT activity with a particulate fraction in the supernatants of infected and transfected cells was verified by subjecting a 50- $\mu$ l sample of an HIV stock to centrifugation in an air-driven ultracentrifuge (28 lb/in<sup>2</sup>, 37 min) (Beckman). An RT assay on 10  $\mu$ l of the uncentrifuged HIV sample indicated 6,210 cpm of activity. A 10- $\mu$ l sample of the centrifuged preparation contained only 57 cpm of activity.

**Analysis of viral proteins.** Immunoblotting was used to identify HIV proteins in lysates prepared from cells transfected with the mutant p7055 and infectious pNL4-3 plasmids. Approximately  $10^8$  SW480 cells were lysed 48 h

posttransfection in 100  $\mu$ l of buffer containing 50 mM Tris hydrochloride, pH 8.0, 0.1 M NaCl, 5 mM EDTA, 1 mM phenylmethylsulfonyl fluoride, 0.5% 3-(3-cholamidopropyl)-dimethylammonio-1-propane sulfonate, 0.2% deoxycholate, and 0.2% aprotinin. Lysates (15  $\mu$ l) were boiled in an equal volume of sample buffer (18), and proteins were resolved on 3 to 27% gradient sodium dodecyl sulfate (SDS)-polyacrylamide gels, followed by electrophoretic transfer to nitrocellulose membranes. The membranes were incubated at room temperature with sera pooled from AIDS patients for 1.5 h and with <sup>125</sup>I-protein A for 1 h, washed, and visualized by autoradiography.

The secretion of newly synthesized HIV gp120 out of transfected cells and into the medium was examined by immunoprecipitating (28) supernatants from SW480 cells previously exposed to pNL4-3 and p7055 DNAs. Cells transfected with mutant or infectious HIV plasmids as described earlier were washed and incubated for 30 min with serum-free Dulbecco modified Eagle medium lacking both methionine and cysteine at 40 h following transfection. The transfected cells were labeled by adding [<sup>35</sup>S]methionine (125  $\mu$ Ci/ml) and cysteine (125  $\mu$ Ci/ml) for 6 h and then propagated for an additional 12 h following the addition of 10% FCS. Tissue culture supernatants were filtered through 0.22- $\mu$ m filters and lysed in an equal volume of RIPA buffer (50 mM Tris hydrochloride, pH 7.5, 0.15 M NaCl, 1% Triton X-100, 1% deoxycholate, and 0.1% SDS) containing 0.5% aprotinin. The extract was preincubated with 10  $\mu$ l of control human serum at 4°C, cleared twice with 50  $\mu$ l of Pansorbin (Calbiochem), and incubated overnight with 2  $\mu$ l of the indicated sera (control human or AIDS patient RJ1370) at 4°C. Precipitates were incubated for 30 min with 10  $\mu$ l of Pansorbin, collected by centrifugation, and washed twice with RIPA buffer and once with water. The precipitated proteins were solubilized by boiling in sample buffer (18), resolved on 10% SDS-polyacrylamide gels, and visualized by autoradiography.

**CD4-binding assay.** The ability of the 7055 mutant gp120 *env* polypeptide to bind to the cellular CD4 gene product was analyzed as described previously (19a). Immunoprecipitates of gp120-CD4 complexes were electrophoresed on 10% SDS-polyacrylamide gels and visualized by autoradiography.

**Revertant cloning and analysis.** Molecular clones of the revertant virus p7055RI were obtained from unintegrated proviral DNA. Cell culture supernatant from day 45 of the p7055RI coculture experiment (see Fig. 5A) was filtered and used to infect A3.01 cells as described above. Approximately  $20 \times 10^6$  cells were harvested on day 13, when numerous syncytia were observed and unintegrated proviral DNA was isolated (14). The unintegrated DNA preparation was restricted with *EcoRI*, ligated to *EcoRI*-restricted  $\lambda$ WES B arms (20) with T4 DNA ligase, packaged in vitro (37), and plated on *Escherichia coli* LE392 cells. Approximately  $5 \times 10^5$  recombinant phage plaques were screened (5) with the <sup>32</sup>P-labeled pBenn6 (10) DNA clone of the lymphadenopathy-associated virus (LAV) provirus. Phage DNAs were restricted with *EcoRI* and *BamHI*, and the resultant 2.7-kb fragments, containing the amino-terminal 2,267 nucleotides of the *env* gene, were inserted into the p7055 mutant plasmid clone. Progeny virions resulting from transfections with these constructions were checked for infectivity in A3.01 cells. The 2.7-kb *EcoRI-BamHI* fragment from revertant HIV proviruses was also transferred into M13 vectors mp18 and mp19, and DNA sequencing was performed as described below.



**DNA sequencing.** DNA was sequenced by using M13 vectors mp18 and mp19, synthetic oligonucleotide primers (38), and the dideoxy chain termination procedure (32).

**Syncytium formation.** The recombinant plasmid pPEenv7h was used for insertion of the 7055 *env* mutant sequence into vaccinia virus. pPEenv7 was constructed as follows. The 3.5-kb *SacI* fragment from HIV strain HTLV-IIIB, which contained the entire envelope gene, was cloned into M13mp18. By site-directed oligonucleotide mutagenesis, an *EcoRV* site followed by the eucaryotic translational consensus sequence CCACC was inserted upstream of the initiating ATG. Another *EcoRV* site was inserted just after the stop codon, allowing easy subcloning of the *env* gene. This 2.5-kb *EcoRV* fragment was subcloned into the *SmaI* site of pSC11, i.e., downstream of the vaccinia virus P7.5 promoter (7). To create a plasmid containing a unique *HindIII* site within the envelope gene, partial *HindIII* digestion, Klenow fill-in, and religation were performed. The resulting plasmid, pPEenv7h, contained unique *StuI* and *HindIII* sites within the envelope gene. The 1.3-kb *StuI-HindIII* region was removed from pPEenv7h and replaced with the analogous fragments from the wild-type pNL4-3 or p7055 mutant clone. The mutagenized area in the two resulting plasmids, pPEenv4-3 and pPEenv7055, respectively, was sequenced for verification. Recombinant vaccinia viruses VPEenv4-3 and VPEenv7055 were constructed by standard methods (7, 22).

Syncytia assays were performed with the HeLa cell line T4Ψ5, which has been transformed with the T4 receptor gene (23). Cells were infected with vaccinia virus recombinants at a multiplicity of infection of 5, 1, or 0.1. The presence or absence of syncytia was determined 16 to 24 h after infection.

## RESULTS

The nucleotide and deduced amino acid sequences of the *env* gene from several different HIV isolates have been determined (3, 9, 27, 30, 35, 39, 41). Computer-assisted analyses of these data have indicated that of eight potential N-linked glycosylation sites present in the gp120 of every isolate, four are clustered within the second highly conserved domain (41), which spans 103 residues and is located about 174 codons from the amino terminus of the processed external envelope protein. We therefore decided to examine the functional significance of this domain by initially targeting the four potential glycosylation sites by in vitro mutagenesis techniques.

**Construction and characterization of HIV *env* mutants.** We have previously described an infectious molecular clone of HIV (pNL4-3) that was constructed by joining two different integrated proviral DNAs at a common *EcoRI* restriction site located at nucleotide position 5779 (1). A 2.7-kb *EcoRI-BamHI* fragment from pNL4-3 DNA (Fig. 1), containing the amino-terminal 2,267 nucleotides of the HIV *env* gene, was subcloned into M13 vectors mp18 and mp19, and specific nucleotide changes were introduced by oligonucleotide-directed mutagenesis (42). Potential N-linked glycosylation sites (Asn-X-Ser/Thr, where X is any amino acid) 6992, 7055, 7099, and 7136 (Fig. 1) (numbers refer to the first nucleotide of the asparagine codon, based on the GenBank proviral nucleotide sequence of the LAV isolate [39]) were altered by changing the asparagine codon. At positions 6992, 7055, and 7136, the asparagine was converted to a glutamine by simultaneously introducing first and third base changes of A to C and T to A, respectively. At site 7099, a single base

change of T to A at the third position resulted in the substitution of lysine for asparagine. The 2.7-kb *EcoRI-BamHI* fragments containing individual amino acid codon changes were then inserted back into the infectious molecular clone, generating four different mutant HIV plasmid clones, p6992, p7055, p7099, and p7136.

Two different biological assays were used for the characterization of these HIV *env* mutants. Virus infectivity was monitored by electroporating the mutant HIV proviral clones into the T4<sup>+</sup> lymphocyte cell line A3.01, which has been shown to be sensitive to viral infection and transfection with HIV proviral DNA (1, 10). The appearance of virion-associated RT activity is indicative of a spreading HIV infection in this T4<sup>+</sup> lymphocyte line. The second assay examined the production of virus particles following calcium phosphate-mediated transfection of the colon carcinoma cell line SW480 (ATCC CCL228), which has been shown to produce virus 24 h after exposure to cloned HIV proviral DNA (1). Although SW480 cells can be infected with HIV, this is an extremely inefficient system compared with T4<sup>+</sup> lymphocytes, since a spreading viral infection cannot be detected by RT assays (2). Consequently, transfection of SW480 cells was done to monitor virus production rather than virus infectivity. In some experiments, the transfected SW480 cells were cocultivated with the T4<sup>+</sup> A3.01 cells 3 days after transfection, permitting assessment of both production and infectivity of progeny virions.

When HIV *env* mutant clones p7099 and p7136 were evaluated for infectivity by electroporation into A3.01 cells, RT activity was detected in the culture supernatants 7 days later and persisted through day 16 (Fig. 2A). The synthesis of RT paralleled that in cultures exposed to the parental infectious pNL4-3 plasmid (Fig. 2A) and was accompanied by syncytium formation and cell death. The infectivity of *env* mutant 6992 was evaluated by transfection of SW480 cells, followed by cocultivation with the T4<sup>+</sup> A3.01 cells as described in the preceding paragraph. The 6922 mutant virus generated a spreading infection in the A3.01 cells that was indistinguishable from that of the infectious pNL4-3 plasmid (data not shown). Thus, elimination of the asparagine codons at sites 6992, 7099, and 7136 in the *env* gene of HIV did not alter infectivity or cytopathic properties.

Unlike the other *env* mutants, 7055 failed to elicit RT activity (Fig. 2A) or generate syncytia in the electroporated A3.01 cells. In other electroporation experiments, this mutant did not produce a spreading infection in the T4<sup>+</sup> lymphocyte line even after more than 60 days (data not shown).

**The 7055 *env* mutation does not affect virion assembly or the synthesis of *env* proteins.** The failure of the 7055 HIV *env* mutant to successfully infect A3.01 cells could reflect a defect in any one of a number of steps (e.g., virion assembly, protein synthesis) culminating in the generation of defective particles. Virion production was monitored by transfecting the p7055 mutant clone into the SW480 colon carcinoma cells. As shown in Fig. 2B, virion-associated RT activity was detected 24, 48, and 72 h following transfection with both p7055 and the infectious pNL4-3 plasmids; however, secondary passage of HIV to the A3.01 lymphocyte line following cocultivation was only observed in cells transfected with the pNL4-3 proviral DNA (Fig. 2B). This again confirmed the inability of the 7055 *env* mutant to initiate a spreading infection in T4<sup>+</sup> cells. A similar result was obtained when RT-positive, cell-free filtrates from p7055- and pNL4-3-transfected colon cells were used for direct infections of A3.01 cells. A spreading viral infection was only detected in



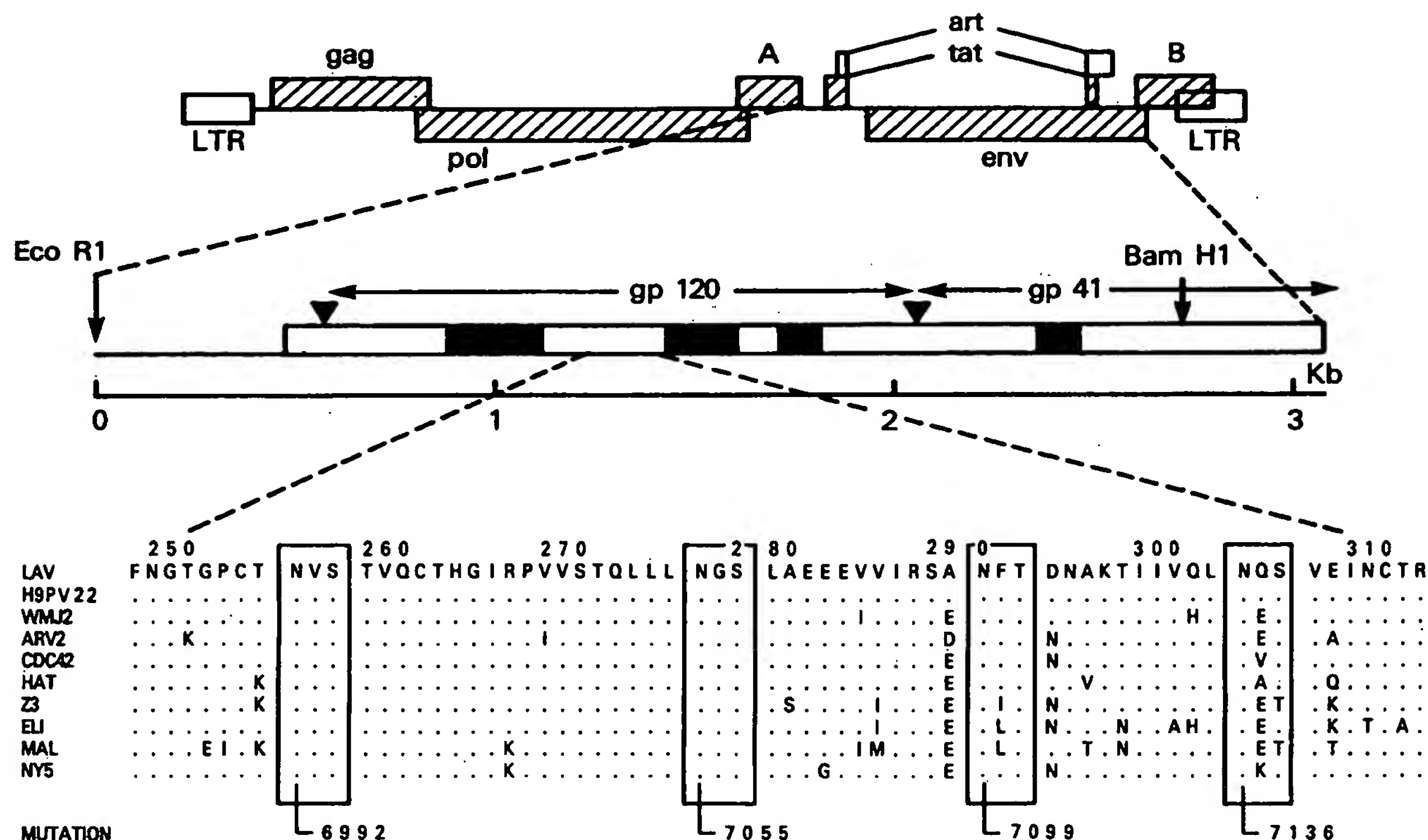


FIG. 1. Locations of potential N-linked glycosylation sites targeted for in vitro mutagenesis. The positions of HIV genes and open reading frames are shown at the top. LTR, Long terminal repeat. Diagrammed in the center is a blow-up of an HIV segment extending from the *Eco*R1 site at 5779 to 8842, the end of the *env* reading frame. Conserved (open) and variable (solid) regions within the *env* gene are indicated. Triangles denote the positions of processing sites. The deduced amino acid sequences of 10 different isolates of HIV (LAV [39], H9PV22 [27], CDC42 [9], Z3 and NY5 [41], and Eli and Mal [3]) from within a portion of the second conserved domain of *gp*120 are shown at the bottom. The numbers indicate amino acid positions. Boxes delineate conserved potential N-linked glycosylation sites (Asn-X-Ser/Thr) in which the asparagine codon was changed. Numbers at the bottom refer to the first nucleotide of each asparagine, as described in the text.

cultures infected with pNL4-3 filtrates during a 37-day observation period (data not shown).

Although the results of the virus production assays indicated that the 7055 *env* mutant directed the synthesis of noninfectious particles possessing RT activity, they did not rule out the possibility that partially assembled virions or only nucleocapsid structures were generated due to an intrinsic defect in *env* protein production. The intracellular synthesis of 7055 viral proteins was examined by immunoblotting and comparing the lysates of SW480 cells transfected with the p7055 mutant and the infectious pNL4-3 plasmid. The protein banding patterns in the two lysates were indistinguishable (Fig. 3A). The presence of *gp*120 and *gp*41 *env* proteins in the p7055-transfected cells indicated that the mutation did not effect the synthesis, processing, or stability of these HIV gene products.

Secretion of 7055 *env* *gp*120 into the culture medium was evaluated by labeling transfected SW480 cells with [ $^{35}$ S] methionine and cysteine, preparing cell-free filtrates of tissue culture supernatants, and immunoprecipitating such preparations with sera from AIDS patients known to have high antibody titers against *gp*120. The two antisera immunoprecipitated *gp*120 from the filtrates of p7055- and pNL4-3-transfected cells (Fig. 3B). In addition, antiserum from AIDS patient RJ1370 precipitated the viral core protein p24. This result, in conjunction with the presence of virion-

associated RT activity in the supernatant (Fig. 2B), suggested that the 7055 mutation did not affect the assembly and release of HIV particles.

To ensure that no additional nucleotide changes had occurred during the generation of the 7055 mutant, reciprocal constructions were made between the p7055 and the pNL4-3 clones by exchanging the 2.7-kb *Eco*R1-*Bam*HI fragment previously used to generate the 7055 mutant (Fig. 1). Both constructs were assayed for infectivity by colon cell transfection, cocultivation with A3.01 cells, and the RT assay. No evidence of a spreading infection was detected with the pNL4-3 clone containing the 2.7-kb *Eco*R1-*Bam*HI fragment of p7055, while the infectivity of p7055 containing the analogous fragment from pNL4-3 was restored (data not shown). Final verification that the asparagine-to-glutamine substitution at position 7055 was solely responsible for the mutant phenotype was obtained by sequencing the entire 2.7-kb *Eco*R1-*Bam*HI fragment of the p7055 plasmid. These results showed that the two introduced nucleotide changes in p7055 were the only differences between it and the parental pNL4-3 clone.

The 7055 mutant *gp*120 polypeptide binds to the CD4 receptor. To determine whether the lack of infectivity associated with the 7055 *env* mutant was due to the inability of its encoded *gp*120 *env* glycoprotein to bind to the CD4 gene product, binding activity was evaluated in vitro with a

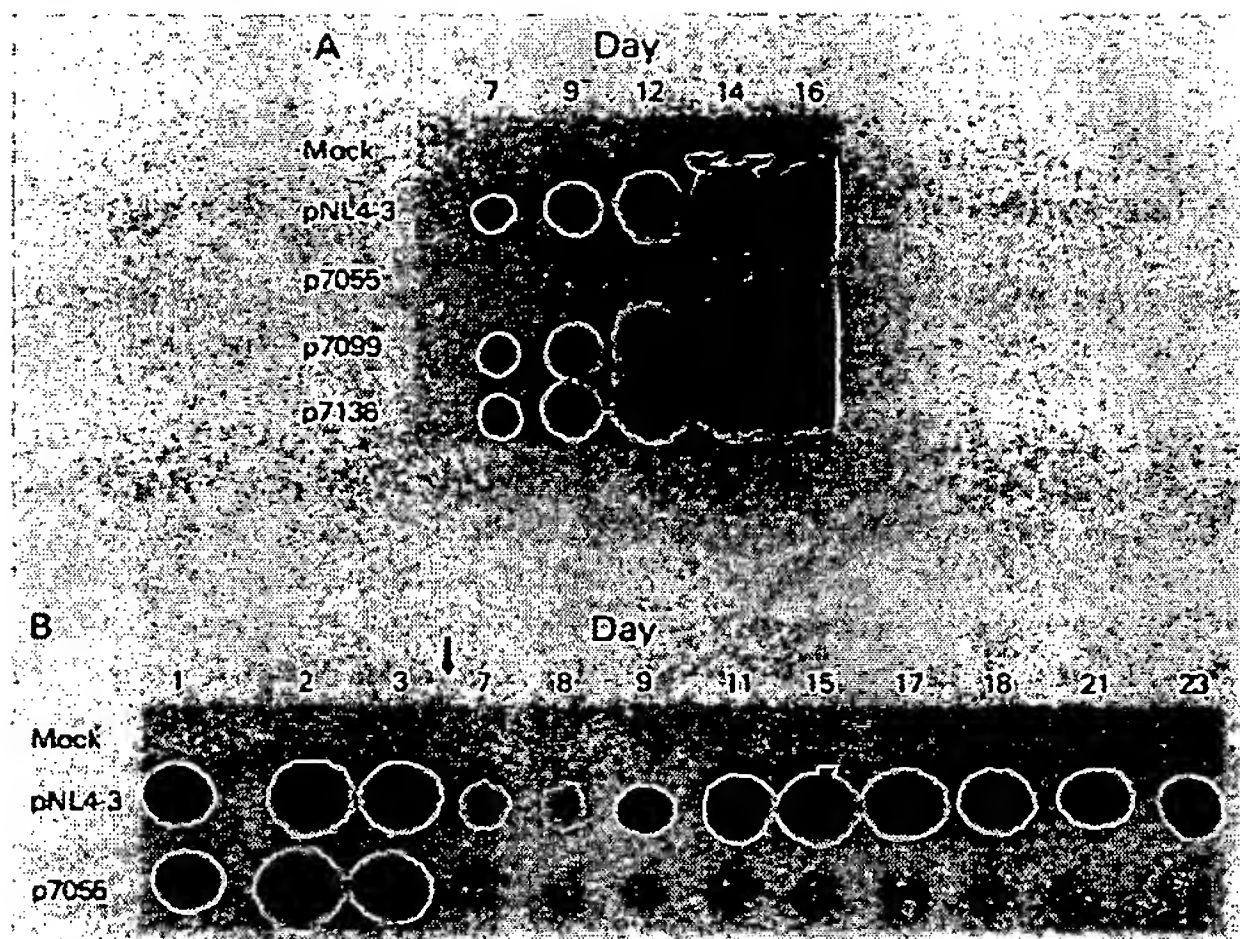


FIG. 2. Infectivity of mutant HIV proviral DNAs. (A) RT activity was measured in 10- $\mu$ l supernatant samples from A3.01 cells electroporated with the infectious (pNL4-3) or *env* mutant (p7055, p7099, and p7136) plasmids at the indicated times as described in the text. (B) RT activity detected in 10- $\mu$ l supernatant samples following calcium phosphate-mediated transfection of SW480 cells with pNL4-3 and p7055 plasmid DNAs. A3.01 cells were cocultured with the transfected SW480 cells beginning on day 3 (as indicated by the arrow).

coimmunoprecipitation assay to measure the formation of the gp120-CD4 molecular complex (24). The 7055 *env* mutation was initially transferred into a gp120 expression vector (19a), which directs the synthesis of a truncated gp120 (gp120-trunc) polypeptide. After transfection of mammalian cells, gp120-trunc is secreted into the medium because it lacks gp41 sequences which anchor the gp120-gp41 envelope protein complex into the plasma membrane. Following transfection of cells with the 7055 gp120-trunc vector, significant levels of [ $^{35}$ S]methionine-labeled polypeptide were detected by immunoprecipitation with gp120-reactive serum from an AIDS patient (Fig. 4). The electrophoretic mobility and heterogeneity were indistinguishable from those of the wild-type gp120-trunc protein.

The binding of 7055 gp120-trunc to CD4 was assessed by incubating transfected-cell supernatants containing the [ $^{35}$ S]methionine-labeled mutant or wild-type protein with detergent-solubilized CD4 receptor prepared from a Chinese hamster ovary (CHO) cell line stably transfected with a human CD4 cDNA expression vector (D. Smith and D. Capon, unpublished results), followed by immunoprecipitation with monoclonal anti-CD4 antibodies to detect the gp120-CD4 complex. Incubation with the OKT4 monoclonal antibody resulted in precipitation of comparable amounts of both the wild-type and mutant gp120-trunc polypeptides similar to that observed with AIDS serum (Fig. 4). Neither gp120-trunc protein was efficiently precipitated with the OKT4A monoclonal antibody, consistent with the behavior of HIV virion-associated gp120 (24), suggesting that the interaction of CD4 with the 7055 *env* mutant gp120, like the wild-type protein, hinders binding of the OKT4A antibody to the CD4 receptor. These results thus demonstrate that the ability to bind CD4 was not grossly impaired by the 7055 *env* mutation.

The 7055 mutant gp120 induces syncytia. The induction of syncytia by HIV has been correlated with the expression of

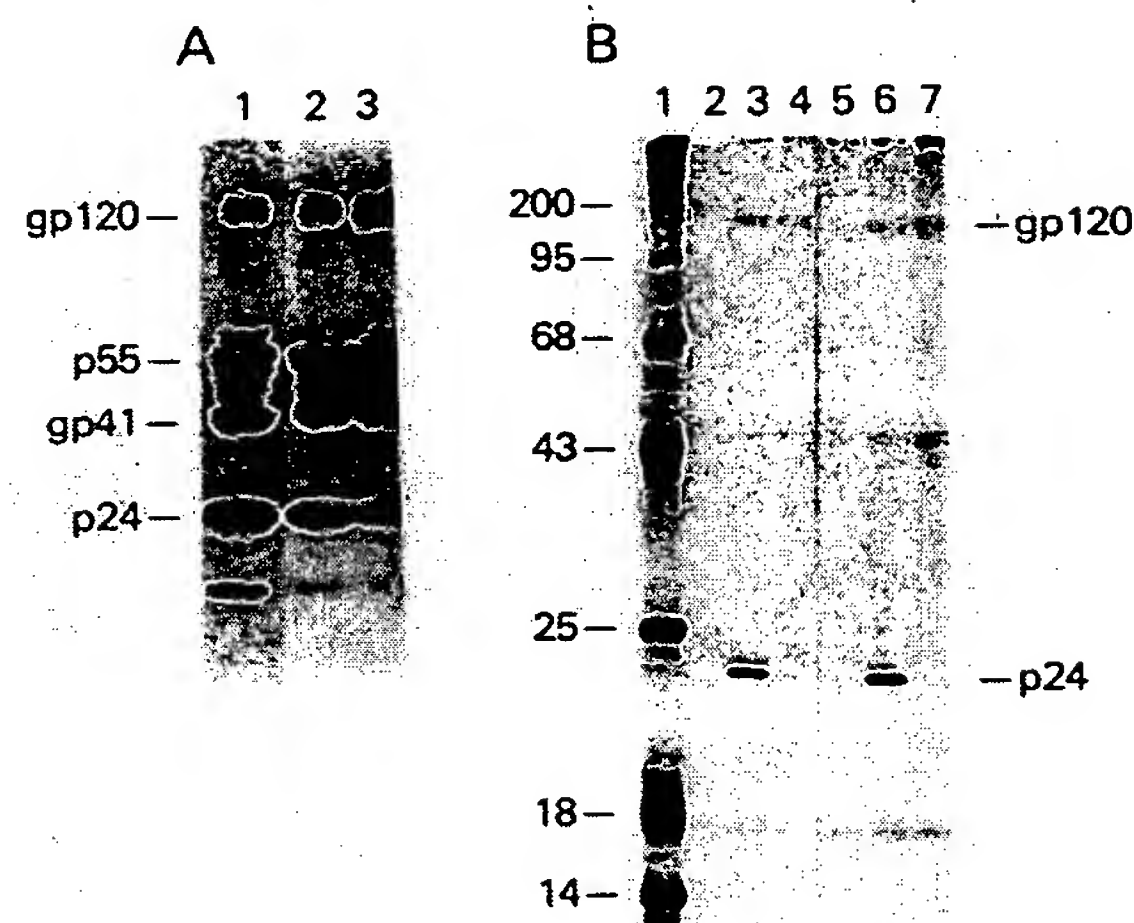


FIG. 3. Production of virus-encoded proteins following transfection of the infectious pNL4-3 and p7055 *env* mutant HIV clones. (A) Immunoblots of lysates from LAV-infected A3.01 cells (lane 1) and pNL4-3 (lane 2) and p7055 (lane 3) lysates from transfected SW480 cells. (B) Immunoprecipitation of  $^{35}$ S-labeled filtrates prepared from SW480 cells transfected with pNL4-3 (lanes 2 to 4) or p7055 (lanes 5 to 7) cloned proviral DNAs. Serum from AIDS patient RJ1370 was used in the immunoprecipitations shown in lanes 3 and 6; serum from AIDS patient RJ97 was used in lanes 4 and 7. Serum from an HIV antibody-negative individual was used in lanes 2 and 5; lane 1 contains molecular weight (in thousands) standards. The positions of the gp120 *env* and p24 *gag* proteins are indicated.

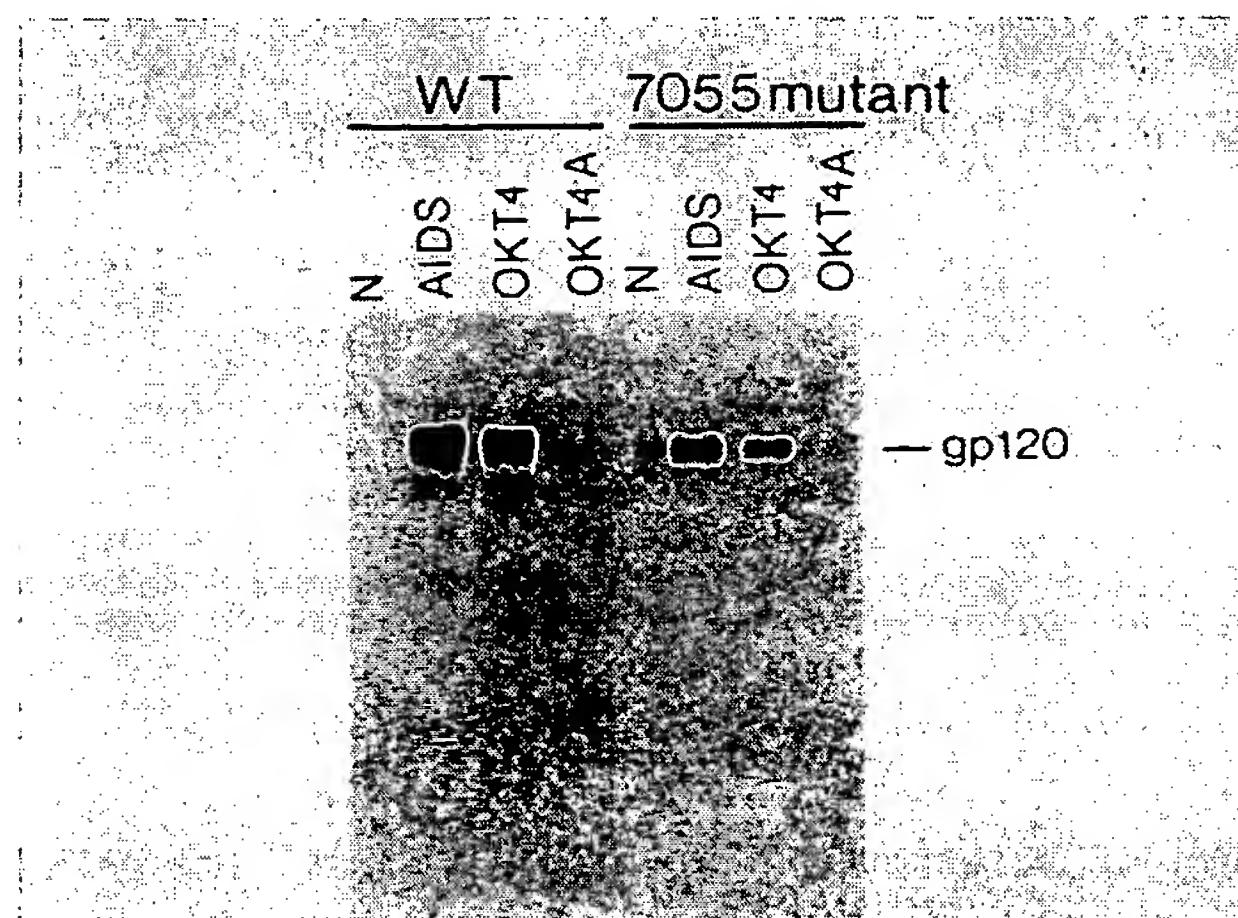


FIG. 4. Coimmunoprecipitation analysis of wild-type *env* and 7055 *env* mutant polypeptide binding to CD4. Wild-type (WT, left) and 7055 mutant (right) gp120-trunc polypeptides were labeled with [ $^{35}$ S]methionine at 60 h following transfection of cells with gp120-trunc expression plasmids. Cell supernatants containing the secreted gp120 proteins were incubated with detergent lysates of CHO CD4 $^{+}$  cells containing solubilized CD4 receptor for 1 h at 4°C to allow the association of gp120 and CD4. Equal portions of the reaction mixture were then analyzed by immunoprecipitation with the following antisera: control human serum L (N), AIDS patient RJ1370 (AIDS) (kindly provided by J. Groopman), OKT4 monoclonal antibody (Ortho), or OKT4A monoclonal antibody (Ortho). The position of gp120-trunc is indicated at the right.



TABLE 1. Infectivity of 7055 region point mutants

Designation <sup>a</sup>	Amino acid sequence <sup>b</sup>	Syncytium formation <sup>c</sup>	Virion-associated RT <sup>c</sup>
pNL4-3 (wild type)	LLNGSL	+	+
p7055	LLQGSL	—	—
p7055D	LL—GSL	—	—
p7052	LQNGSL	—	—
p7058	LLNDL	—	—

<sup>a</sup> Numbers in mutant designations refer to the first nucleotide of the mutated codon based on the GenBank proviral nucleotide sequence of the LAV isolate (38).

<sup>b</sup> Amino acid changes are underlined; deletion is shown with a dash.

<sup>c</sup> Data are from infection of A3.01 cells during a 37-day period for pNL4-3, p7055, p7052, and p7058; p7055D results were obtained from cocultures over a 34-day period.

CD4 on the surface of lymphocytes (34) and HeLa cells (23). Since the functional defect resulting from the 7055 mutation did not appear to affect CD4 binding, the ability of the mutant *env* gp120 to form syncytia was assayed with recombinant vaccinia virus (22) as described in Materials and Methods. The HeLa cell line T4Ψ5 was infected with VPEenv4-3, VPEenv7055, VPEenv7, and wild-type vaccinia virus. Similar numbers of giant syncytia were observed in cultures infected with recombinant vaccinia constructions 16 to 24 h postinfection (data not shown). Cells infected with wild-type vaccinia virus did not form syncytia.

**Additional point mutations in the 7055 region of the HIV *env* gene.** The defective phenotype resulting from the position 7055 substitution suggested that the asparagine codon was critical for HIV infectivity. To explore this possibility further, oligonucleotide-directed mutagenesis was used to delete this codon. No evidence of a spreading viral infection was detected with progeny virions containing this deletion (Table 1). Since this codon lies within a highly conserved region of the HIV *env* gene, we next wanted to determine whether a larger domain which encompassed the 7055 site might be involved in viral infectivity. Nucleotide changes were therefore introduced into the two codons adjacent to Asn-7055 (Table 1). The resulting two mutant proviral DNAs (7052 and 7058) also failed to elicit a spreading infection (Table 1). Immunoblotting of transfected SW480 cells indicated that these mutations did not affect the synthesis and processing of HIV proteins (data not shown).

**Reversion of the 7055 HIV *env* mutant in tissue culture.** During the course of several long-term cocultures, set up to monitor marker rescue of the 7055 mutant, the appearance of infectious virions was noted in some of the p7055 DNA controls. Of nine extended SW480 transfections–A3.01 cocultures with p7055 DNA, three resulted in particle-associated RT activity 26 to 35 days following the introduction of the mutant proviral DNA. Since RT activity was consistently present within 5 days in cocultures of the pNL4-3 infectious clone, the delayed appearance of RT suggested that 7055 revertants had emerged. RT activity was not detected during a 45-day observation period in the six negative cocultures. Selected samples and time points from the nine experiments are shown in Fig. 5A.

To assess directly the infectivity of one of the putative revertants, supernatants from p7055R1 (day 45, Fig. 5A) and pNL4-3 (day 12, Fig. 5A) transfection-cocultures were used to establish infections of the A3.01 lymphocytes as described in Materials and Methods. Each inoculum was normalized for particle number by using comparable amounts of RT activity. The p7055R1 virus exhibited RT activity (Fig. 5B),

syncytium formation, and killing of the A3.01 cells, although the onset of these activities was delayed relative to the pNL4-3 infection.

The emergence of infectious virus from the long-term cocultures of the p7055 *env* mutant suggested that a reversion in the *env* gene had occurred. To verify the molecular basis for such a reversion, A3.01 cells were infected with filtrates from 7055R1-producing cells, and cloned unintegrated proviral DNA was obtained by using a λWES vector, as described in Materials and Methods. Analytic restriction enzyme digestion of two lambdaphage clones revealed the presence of the 2.7-kb *EcoRI*–*Bam*HI fragment containing HIV *env* gene sequences. These DNA fragments were introduced into the p7055 mutant clone, and the infectivity of the reconstructed proviral DNAs was examined. In both instances, replacement of the mutant *EcoRI*–*Bam*HI fragment with the analog from the revertant restored infectivity, indicating that the change(s) responsible for the revertant phenotype was present in this region of the viral genome.

The complete nucleotide sequence of the 2.7-kb *EcoRI*–*Bam*HI fragments from both revertant clones was determined following their subcloning into M13 vectors mp18 and mp19. Sequencing of the revertant *env* segments demonstrated the preservation of the original asparagine-to-glutamine mutation at position 7055 in both cases; however, new nucleotide changes were identified (Fig. 6) in each

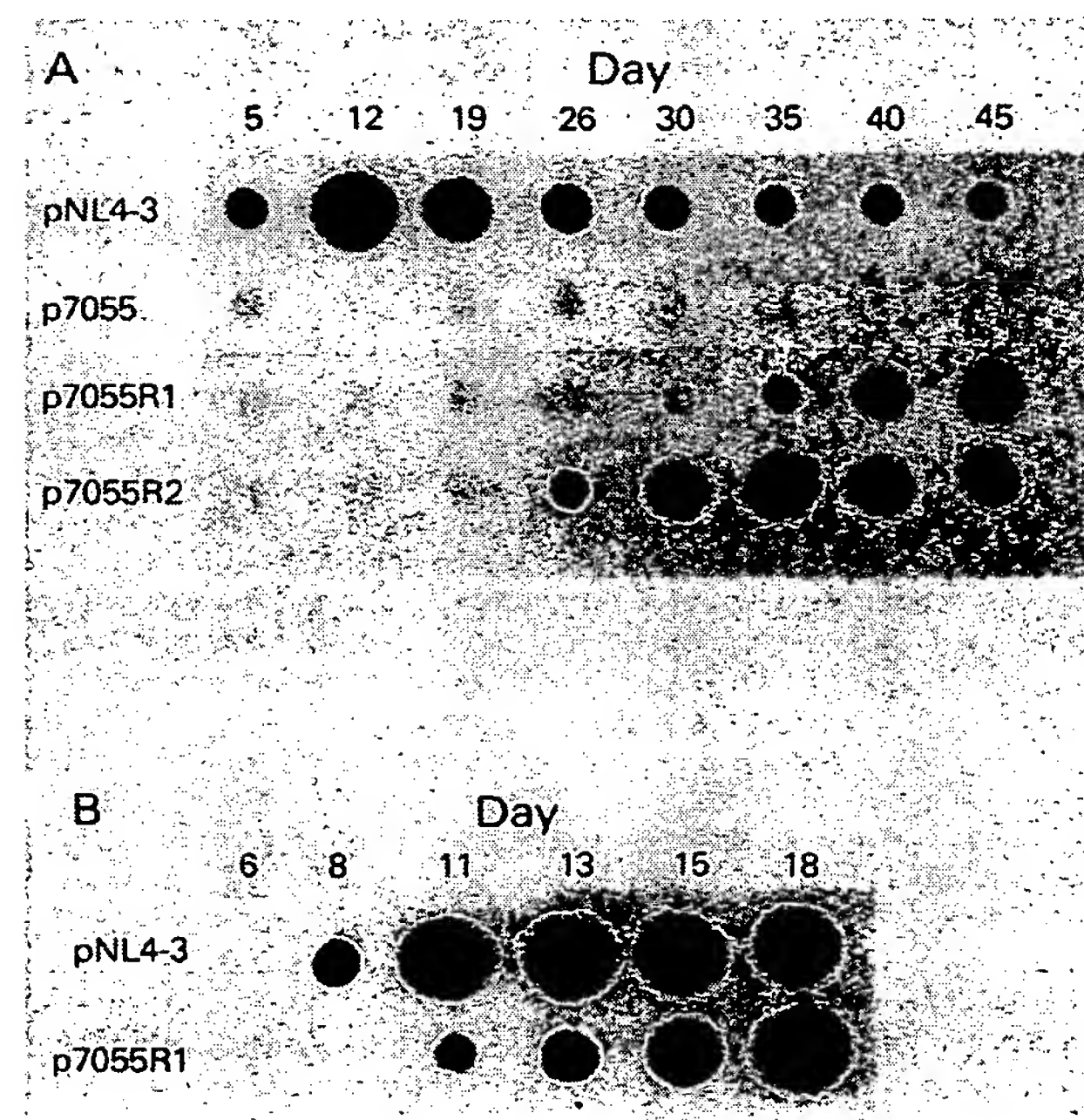


FIG. 5. Detection and characterization of revertant virus from long-term cocultures of the 7055 mutant. (A) SW480 cells were transfected with pNL4-3 or p7055 plasmid DNA; A3.01 cells ( $2 \times 10^6$ ) were added 24 h later, and RT assays were carried out on the culture supernatants at the indicated times. The infectious virions detected in p7055R1 and p7055R2 represent the results of two experiments among nine conducted in which revertants appeared in cocultures of the p7055 mutant. (B) RT activities in supernatants of A3.01 cells infected with HIV prepared from pNL4-3 (day 12, panel A) or p7055R1 (day 45, panel A) as described in the text.



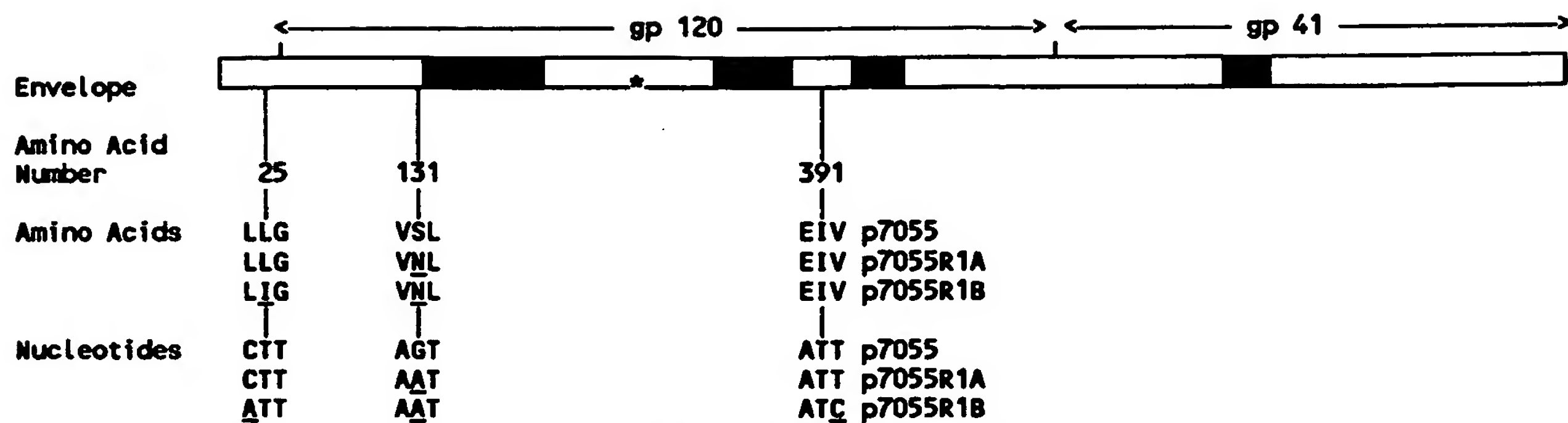


FIG. 6. Nucleotide and amino acid substitutions identified in the revertant proviral clones. Conserved and variable portions of the HIV envelope are diagrammed as indicated in the legend to Fig. 1. The nucleotide and amino acid substitutions identified in the revertant proviral clones p7055R1A and p7055R1B are shown. The asterisk indicates the original asparagine-to-glutamine substitution at position 7055 which was also present in both revertant clones.

clone. One clone (p7055R1A) contained a single nucleotide substitution (a G to A change) in codon 131, leading to the substitution of an asparagine for a serine. The second clone (p7055R1B) contained three nucleotide changes, all of which mapped to the gp120 coding region. The first, located within the leader sequence at codon 25, involved substitution of an A for a C in the first position and resulted in a codon change from leucine to isoleucine. The second was identical to the single change in the first revertant clone at codon 131 and caused the substitution of an asparagine for a serine. The final nucleotide change occurred at codon 391 and was a third-position base change of T to C, which did not alter the codon specifying isoleucine.

#### DISCUSSION

Although the *env* genes of different HIV isolates exhibit profound heterogeneity at both the nucleotide and deduced amino acid levels, several domains within both gp120 and gp41 are highly conserved, suggesting that they mediate properties or functions common to all isolates. For example, a 102-amino-acid polypeptide, between residues 28 and 128 of gp41, specifically reacts with sera from HIV antibody-positive individuals in enzyme-linked and immunoblotting assays (6). A more recent study (12) done with a series of synthetic oligopeptides localized an immunodominant domain within a conserved portion of gp41 (codons 81 to 92) that is recognized by 100% of the sera from HIV-infected persons. Another property of all HIV isolates is the interaction between gp120 and its cellular receptor, CD4. Using murine monoclonal antibodies and oligonucleotide-directed mutagenesis, Laskey et al. (19a) identified an epitope within the fourth conserved domain of gp120 (41) that was involved in the binding to CD4. Deletions and a point mutation within this region (amino acids 397 to 439) resulted in significantly reduced binding. A feature of HIV infection of CD4<sup>+</sup> cells, thought to involve *env* proteins, is the formation of syncytia that commonly precedes the peak of RT activity. A mutant *env* gene containing three hydrophilic amino acid substitutions within the conserved hydrophobic N-terminus of gp41 failed to induce syncytia following its introduction into CD4<sup>+</sup> HeLa cells (R. Willey, J. Felser, and J. Silver, unpublished observation), defining still another functionally important region of the HIV envelope.

At present, the nature of both the structural and functional defect in the 7055 mutant is unknown. The results described in this paper indicate that the binding of the 7055 mutant gp120 to the CD4 protein was indistinguishable from that of wild-type gp120. Although the interaction of the mutant *env* protein with CD4 was not analyzed in competition experiments with a large excess of unlabeled mutant gp120, the validity of the binding results is supported by two additional pieces of data. First, vaccinia virus recombinants expressing wild-type or 7055 mutant gp120 elicited similar amounts of giant syncytia following infection of CD4<sup>+</sup> HeLa cells. Since syncytium formation depends on the interaction of HIV envelope-expressing cells with cells containing CD4 protein on their surface (34), it must be concluded that the 7055 mutant *env* protein binds to CD4 in a functionally relevant manner. Furthermore, two murine monoclonal antibodies to the HIV envelope have been obtained which react with epitopes mapping in the second conserved domain of gp120 (G. Nakamura and L. Lasky, manuscript in preparation). One of these binds to the very region of gp120 encompassing the 7055 mutation. Significantly, neither of the monoclonal antibodies blocks the binding of wild-type gp120 to CD4.

The inability to initiate a productive infection by eliminating the highly conserved potential N-linked glycosylation site at 7055 points to a critical role for an attached sugar group at this position of gp120; however, no data presently exist demonstrating that the 7055 site is actually glycosylated in the mature gp120. The results presented in Table 1, showing that mutations in codons adjacent to 7055 also rendered the virus noninfectious, suggest the existence of a domain (rather than a critical glycosylation site) within gp120 which is involved in an early step of viral infection. In this regard, the 11 residues surrounding the 7055 site are invariant among the 10 HIV isolates shown in Fig. 1. Our results suggest that the initiation of an HIV infection of CD4<sup>+</sup> cells may involve more than binding of the virus to its receptor. Subsequent modification(s) of the viral envelope, fusion of the viral and cellular membranes (36), penetration, endocytosis, and partial uncoating represent only some of the steps that must occur following adsorption. One or more of these steps may require an intact (or functionally equivalent) 7055 domain.

An unexpected result from these studies was the identification of second-site revertants of the 7055 *env* mutant

which occurred in tissue culture. Sequencing analyses of the two revertant *env* genes point to the codon 131 substitution as the functionally significant change. It is not obvious how the codon 131 alteration might compensate for the original 7055 mutation. One possibility is that the introduction of an asparagine codon at this position provides a novel site for N-linked glycosylation, but this seems very unlikely since this change did not generate a canonical site. All reported sequences of HIV *env* genes contain either a serine or threonine at codon 131. The elimination of this residue might prevent O-linked glycosylation. Since serine and threonine can also serve as targets of phosphorylation, the substitution of asparagine at this position would preclude such a modification and its potential effect on gp120. The effect of other changes at codon 131 of gp120 are currently being evaluated.

It is not presently clear in which type of cell (colon or lymphocyte) the reversions occurred. Several facts relevant to detection of the revertants should be noted. First, SW480 colon carcinoma cells are capable of producing infectious HIV for longer than 60 days following transfection with pNL4-3 DNA when cultured in the absence of T4<sup>+</sup> lymphocytes (K. Strebel and M. Martin, unpublished observation). This constitutive production of progeny particles most likely reflects the intrinsic sensitivity of SW480 cells to transfection (1) coupled with the establishment of a low level of chronic HIV infection of this cell line (2). Second, direct introduction of the 7055 proviral DNA into A3.01 cells by electroporation consistently failed to generate infectious virions even after 60 days. The cocultivation of virus-producing SW480 cells with susceptible A3.01 cells, however, may have allowed the "selection" and amplification of the revertant infectious particles. Since the original 7055 mutation involved nucleotide substitutions rather than multiple codon deletions within the *env* gene, it is conceivable that the defect is conditional rather than absolute. In a milieu containing continuously produced 7055 mutant virus particles, A3.01 cells may be inefficiently yet definitively infected, thereby establishing a replicative cycle which would include an obligate reverse transcription step that gave rise to revertant particles in cocultures.

The results from this study have provided additional information about the role of the HIV envelope during viral replication. The interaction of envelope with T lymphocytes involves more than binding of virions to the CD4 receptor. Subsequent events associated with internalization of virus particles, including penetration and endocytosis, may be modulated by the portion of gp120 defined by the 7055 mutant. These findings raise the possibility that interfering HIV mutants or *env* gene products encompassing this segment might compete with replication-competent virus, as described for the FV-4 resistance determinant (16) in the murine leukemia virus system. In addition, antibodies directed against this region of the HIV envelope might be useful in blocking virus infection.

Although the region encompassing the 7055 site appears to be critical for HIV infectivity, specific residues within this area may not be absolutely required. The emergence of the second-site revertant virus indicates that a change in a different portion of gp120 can partially compensate for the original 7055 mutation. The appearance of the revertant may involve a mechanism(s) that allows the survival and persistence of HIV. This could be especially important in the presence of an active immune response if viral progeny arose that were able to escape neutralization. Further studies into the potential significance of variability during HIV-induced disease are presently in progress.

#### ACKNOWLEDGMENTS

We thank Klaus Strebel and Kathleen Clouse for performing the immunoblotting. We also thank Gilbert Ashwell, Peter McPhie, and Eugenio Santos for helpful discussions concerning protein structure and function.

#### LITERATURE CITED

1. Adachi, A., H. E. Gendelman, S. Koenig, T. Folks, R. Willey, A. Rabson, and M. A. Martin. 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and non-human cells transfected with an infectious molecular clone. *J. Virol.* 59:284-291.
2. Adachi, A., S. Koenig, H. E. Gendelman, D. Daugherty, S. Gattoni-Celli, A. S. Fauci, and M. A. Martin. 1987. Productive persistent infection of human colorectal cell lines with human immunodeficiency virus. *J. Virol.* 61:209-213.
3. Alizon, M., S. Wain-Hobson, L. Montagnier, and P. Sonigo. 1986. Genetic variability of the AIDS virus: nucleotide sequence analysis of two isolates from African patients. *Cell* 46:63-74.
4. Barré-Sinoussi, F., J. C. Cherman, R. Rey, M. T. Nugeryre, S. Chamaret, J. Gruest, C. Dauguet, C. Axler-Blin, F. Brun-Vezinet, C. Rouzioux, W. Rosenbaum, and L. Montagnier. 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 220:868-871.
5. Benton, D., and R. W. Davis. 1977. Screening  $\lambda$ gt recombinant clones by hybridization to single plaques in situ. *Science* 196:180-182.
6. Cabradilla, C. D., J. E. Groopman, J. Lanigan, M. Renz, L. A. Lasky, and D. J. Capon. 1986. Serodiagnosis of antibodies to the human AIDS retrovirus with a bacterially synthesized *env* polypeptide. *Bio/Technology* 4:128-133.
7. Chakrabarti, S., K. Brechling, and B. Moss. 1985. Vaccinia virus expression vector: coexpression of  $\beta$ -galactosidase provides visual screening of recombinant virus plaques. *Mol. Cell. Biol.* 5:3403-3409.
8. Chopin, P. W., and A. Scheld. 1980. The role of viral glycoproteins in adsorption, penetration, and pathogenicity of viruses. *Rev. Infect. Dis.* 2:40-61.
9. Desai, S. M., V. S. Kalyanaraman, J. M. Casey, A. Srinivasan, P. R. Andersen, and S. G. Devare. 1986. Molecular cloning and primary nucleotide sequence analysis of a distinct human immunodeficiency virus isolate reveal significant divergence in its genomic sequences. *Proc. Natl. Acad. Sci. USA* 83:8380-8384.
10. Folks, T., S. Benn, A. Rabson, T. Theodore, M. D. Hoggan, M. Martin, M. Lightfoote, and K. Sell. 1985. Characterization of a continuous T-cell line susceptible to the cytopathic effects of the acquired immunodeficiency syndrome (AIDS)-associated retrovirus. *Proc. Natl. Acad. Sci. USA* 82:4539-4543.
11. Gallo, R. C., S. Z. Salahuddin, M. Popovic, G. M. Schearer, M. Kaplan, B. F. Haynes, T. J. Palker, R. Redfield, J. Oleske, B. Safai, G. White, P. Foster, and P. D. Markham. 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* 224:500-503.
12. Gnann, J. W., Jr., J. A. Nelson, and M. B. A. Oldstone. 1987. Fine mapping of an immunodominant domain in the transmembrane glycoprotein of human immunodeficiency virus. *J. Virol.* 61:2639-2641.
13. Goff, S., P. Traktman, and D. Baltimore. 1981. Isolation and properties of Moloney murine leukemia virus mutants: use of a rapid assay for release of virion reverse transcriptase. *J. Virol.* 38:239-248.
14. Hirt, B. 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.* 26:364-369.
15. Homma, M., and M. Onuchi. 1973. Trypsin action on the growth of Sendai virus in tissue culture cells. *J. Virol.* 12:1457-1465.
16. Ikeda, H., and T. Odaka. 1983. Cellular expression of murine leukemia virus gp-70 related antigen on thymocytes of uninfected mice correlated with Fv-4 gene-controlled resistance to Friend leukemia virus infection. *Virology* 133:65-76.
17. Klatzmann, D., F. Barré-Sinoussi, M. T. Nugeryre, C. Dauguet, E. Vilmer, C. Griscelli, F. Brun-Vezinet, C. Rouzioux, J. C.



- Gluckman, J. C. Chermann, and L. Montagnier. 1984. Selective tropism of lymphadenopathy associated virus (LAV) to helper-inducer T lymphocytes. *Science* 225:59-62.
18. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* 227:680-685.
  19. Lane, H. C., and A. S. Fauci. 1985. Immunologic abnormalities in the acquired immunodeficiency syndrome. *Annu. Rev. Immunol.* 3:477-500.
  - 19a. Lasky, L. A., G. Nakamura, J. H. Smith, C. Fennie, L. Shimosaki, E. Patzer, P. Berman, T. Gregory, and D. J. Capon. 1987. Delineation of a region of the human immunodeficiency virus type 1 (HIV-1) gp120 glycoprotein critical for interaction with the CD4 receptor. *Cell* 50:975-985.
  20. Leder, P., D. Tiemeier, and L. Enquist. 1977. Ek2 derivatives of bacteriophage lambda useful in the cloning of DNA from higher organisms: the  $\lambda$ gt WES system. *Science* 196:175-177.
  21. Levy, J. A., A. D. Hoffman, S. M. Kramer, J. A. Lanois, J. M. Shimabukuro, and L. S. Oskiro. 1984. Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science* 225:840-842.
  22. Mackett, M., G. L. Smith, and B. Moss. 1984. General method for production and selection of infectious vaccinia virus recombinants expressing foreign genes. *J. Virol.* 49:857-864.
  23. Maddon, P. J., A. G. Dalgleish, J. S. McDougal, P. R. Clapham, R. A. Weiss, and R. Axel. 1986. The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell* 47:333-348.
  24. McDougal, J. S., M. S. Kennedy, J. N. Sligh, S. P. Cort, A. Mawle, and J. K. A. Nicholson. 1986. Binding of HTLV-III/LAV to T4<sup>+</sup> cells by a complex of the 110K viral protein and the T4 molecule. *Science* 231:382-385.
  25. McDougal, J. S., A. Mawle, S. P. Cort, J. K. A. Nicholson, D. G. Cross, J. A. Scheppeler-Campbell, D. Hicks, and J. Sligh. 1985. Cellular tropism of the human retrovirus HTLV-III/LAV. I. Role of T cell activation and expression of the T4 antigen. *J. Immunol.* 135:3151-3162.
  26. McDougal, J. S., J. K. A. Nicholson, D. G. Cross, S. P. Cort, S. M. Kennedy, and A. C. Mawle. 1986. Binding of the human retrovirus HTLV-III/LAV/ARV/HIV to the CD4 (T4) molecule: conformation dependence, epitope mapping, antibody inhibition, and potential for idiotypic mimicry. *J. Immunol.* 137:2937-2944.
  27. Muesing, M. A., D. H. Smith, C. D. Cabradilla, C. V. Benton, L. A. Lasky, and D. J. Capon. 1985. Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus. *Nature (London)* 313:430-458.
  28. Muesing, M. A., D. H. Smith, and D. J. Capon. 1987. Regulation of mRNA accumulation by a human immunodeficiency virus trans-activator protein. *Cell* 48:691-701.
  29. Potter, H., W. Lawrence, and P. Leder. 1984. Enhancer-dependent expression of human K immunoglobulin genes introduced into mouse pre-B lymphocytes by electroporation. *Proc. Natl. Acad. Sci. USA* 81:7161-7165.
  30. Sanchez-Pescador, R., M. D. Power, P. J. Barr, K. S. Steimer, M. M. Stempien, S. L. Brown-Shimer, W. W. Gee, A. Ranard, A. Randolph, J. A. Levy, D. Dina, and P. A. Luciw. 1985. Nucleotide sequence and expression of the AIDS-associated retrovirus (ARV-2). *Science* 227:484-492.
  31. Sanchez-Pescador, R., and M. S. Urdea. 1984. Use of purified synthetic deoxynucleotide primers for rapid dideoxynucleotide chain termination sequencing. *DNA* 3:339-343.
  32. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
  33. Scheid, A., and P. W. Choppin. 1974. Identification of biological activities of paramyxovirus glycoproteins. Activation of cell fusion, hemolysis, and infectivity by proteolytic cleavage of an inactive precursor protein by Sendai virus. *Virology* 57:475-490.
  34. Sodroski, J., W. C. Goh, C. Rosen, K. Campbell, and W. A. Haseltine. 1986. Role of the HTLV-III/LAV envelope in syncytium formation and cytopathicity. *Nature (London)* 322:470-474.
  35. Starcich, B. R., B. H. Hahn, G. M. Shaw, P. D. McNeely, S. Modrow, H. Wolf, E. S. Parks, W. P. Parks, S. F. Josephs, R. C. Gallo, and F. Wong-Staal. 1986. Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS. *Cell* 45:637-648.
  36. Stein, B. S., S. D. Gowda, J. D. Lifson, R. C. Penhallow, K. G. Bensch, and E. G. Engleman. 1987. pH-independent HIV entry into CD4-positive T cells via virus envelope fusion to the plasma membrane. *Cell* 49:659-668.
  37. Sternberg, N., D. Tiemeier, and L. Enquist. 1977. *In vitro* packaging of a  $\lambda$  Dam vector containing EcoRI DNA fragments of *Escherichia coli* and phage P1. *Gene* 1:255-280.
  38. Strauss, E. C., J. A. Kobori, G. Siu, and L. E. Hood. 1986. Specific-primer-directed DNA sequencing. *Anal. Biochem.* 154:353-360.
  39. Wain-Hobson, S., P. Sonigo, O. Danos, S. Cole, and M. Alizon. 1985. Nucleotide sequence of the AIDS virus, LAV. *Cell* 40:9-17.
  40. Wigler, M., A. Pellicer, S. Silverstein, R. Axel, G. Urtsaub, and L. Chasin. 1979. DNA-mediated transfer of the adenine phosphoribosyl transferase locus into mammalian cells. *Proc. Natl. Acad. Sci. USA* 76:1373-1376.
  41. Willey, R. L., R. A. Rutledge, S. Dias, T. Folkes, T. Theodore, C. E. Buckler, and M. A. Martin. 1986. Identification of conserved and divergent domains within the envelope gene of the acquired immunodeficiency syndrome retrovirus. *Proc. Natl. Acad. Sci. USA* 83:5038-5042.
  42. Zoller, M. J., and M. Smith. 1984. Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template. *DNA* 3:479-488.



## Effects of Deletions in the Cytoplasmic Domain on Biological Functions of Human Immunodeficiency Virus Type 1 Envelope Glycoproteins

DANA H. GABUZDA,<sup>1</sup> ANDREW LEVER,<sup>2†</sup> ERNEST TERWILLIGER,<sup>3</sup> AND JOSEPH SODROSKI<sup>2,3\*</sup>

Division of Human Retrovirology, Dana-Farber Cancer Institute, 44 Binney Street,<sup>2</sup> and Departments of Pathology<sup>3</sup> and Neurology,<sup>1</sup> Harvard Medical School, Boston, Massachusetts 02115

Received 31 December 1991/Accepted 25 February 1992

The role of the cytoplasmic domain of the human immunodeficiency virus type 1 (HIV-1) envelope glycoproteins in virus replication was investigated. Deletion of residues 840 to 856 at the carboxyl terminus of gp41 reduced the efficiency of virus entry during an early step in the virus life cycle between CD4 binding and formation of the DNA provirus without affecting envelope glycoprotein synthesis, processing, or syncytium-forming ability. Deletion of residues amino terminal to residue 846 was associated with decreased stability of envelope glycoproteins made in COS-1 cells, but this phenotype was cell type dependent. The cytoplasmic domain of gp41 was not required for the incorporation of the HIV-1 envelope glycoproteins into virions. These results suggest that the carboxyl terminus of the gp41 cytoplasmic domain plays a role in HIV-1 entry other than receptor binding or membrane fusion. The cytoplasmic domain of gp41 also affects the stability of the envelope glycoprotein in some cell types.

The presence of a long intracytoplasmic domain at the carboxyl terminus of the gp41 transmembrane protein of the human immunodeficiency virus type 1 (HIV-1) distinguishes the HIV-1 envelope glycoprotein from that of most retroviruses (4, 12, 34). The transmembrane proteins of the type C and type D retroviruses terminate within 50 amino acids carboxy terminal to the transmembrane domain (34). The presence of a cytoplasmic domain extending more than 100 amino acids beyond the transmembrane region in HIV-1, the simian immunodeficiency virus (SIV), visna virus, and the equine infectious anemia virus suggests a specific function for this region in the life cycle of the lentivirus subfamily of retroviruses (12, 34). HIV-2 and SIV are frequently truncated just after the membrane anchor domain, but these truncated forms appear to result from selection during tissue culture propagation in human cells (4, 19, 20, 22). The natural form of the SIV transmembrane protein is the full-length 41-kDa protein (20, 22).

The HIV-1 envelope glycoprotein is initially made as a 160-kDa glycosylated precursor that is cleaved to form the gp120 exterior and gp41 transmembrane subunits (7, 8, 40). The gp120 glycoprotein determines the tropism of HIV-1 for specific target cells by binding to the CD4 molecule (6, 21, 27). The gp41 transmembrane glycoprotein contains an extracellular domain that is required for membrane fusion and noncovalent association with the gp120 glycoprotein, a hydrophobic transmembrane region that anchors the gp120-gp41 complex in the cell or virion lipid bilayer, and an intracytoplasmic domain of unknown function (2, 11, 14, 24, 35). The cytoplasmic domain contains two highly conserved segments with the potential to form amphipathic  $\alpha$  helices and may form a secondary association with the lipid bilayer (1, 14, 15, 39).

The cytoplasmic domain of gp41 is required for efficient HIV-1 replication, but the function of this domain in the

virus life cycle is unknown (9, 26, 37). The cytoplasmic domain is not required for syncytium formation or binding to CD4 (7, 8, 24) but may modulate intracellular transport and processing of gp160 in some cell types (16). Whether the cytoplasmic domain is involved in virion assembly, for example, by facilitating the incorporation of envelope glycoproteins into virions, is unknown.

To investigate the role of the gp41 cytoplasmic domain in virus replication, the effect of deletions in this region on the synthesis, processing, cell surface expression, incorporation into virions, and biological function of the HIV-1 envelope glycoproteins was investigated.

### MATERIALS AND METHODS

**Plasmids.** The pSVIIIenv plasmid expresses the HIV-1 *env* and *rev* genes of the HXB2 strain (31) under the control of the HIV-1 long terminal repeat (LTR) (17, 24). The pHXBΔenvCAT plasmid contains an HIV-1 provirus with an in-frame deletion from the *Bgl*II to *Bgl*II sites (nucleotides 6620 and 7200) of the sequence of Ratner et al. (32) in the *env* gene and a chloramphenicol acetyltransferase (CAT) gene replacing the *nef* gene (17, 36). The pHXBΔenvΔEcoCAT plasmid, derived from pHXBΔenvCAT, contains an additional deletion from *Eco*RI to *Eco*RI (nucleotides 4231 to 5325) overlapping the *pol* gene. These plasmids contain a simian virus 40 origin of replication. Envelope glycoprotein deletion mutants were made by *Bal* 31 exonuclease digestion at either the *Bam*HI site or the *Xho*I site (nucleotide 8053 or 8476, respectively) (24, 32, 37). The *env* open reading frame was restored when necessary either by creating blunt ends prior to religation or by inserting *Cla*I linkers. The Δ(767-856) mutant was made by inserting an *Eco*RI linker at an *Mn*II site to create a frameshift at amino acid 767. Oligonucleotide-directed mutagenesis was used to create the Δ(796-803), Δ(796-804), Δ(846-856), Δ(851-856), and Δ(854-856) deletions in an *Eco*RI-to-*Xho*I (nucleotides 5325 to 8476) fragment of HXB2 subcloned into the pBluescript plasmid (Stratagene) (25). All plasmids were sequenced in the region

\* Corresponding author.

† Present address: Department of Medicine, Addenbrooke's Hospital, Cambridge, CB2 2QQ United Kingdom.

of the introduced mutation. The pHXB2Δ(840-856) provirus was made by subcloning the deleted *Bam*HI-to-*Xho*I fragment into the pHXB2 plasmid (31). The *rev* expressor plasmid pSVIIIenvΔKS contains a *Kpn*I-to-*Stu*I (nucleotide 5926 to 6411) out-of-frame deletion in the envelope gene and expresses HIV-1 *rev* but not the envelope glycoproteins.

**Virus replication studies in Jurkat lymphocytes.** Jurkat T lymphocytes were transfected by the DEAE-dextran method (30) with 10 μg of pHXB2 or pHXB2Δ(840-856). Following transfection, cultures were maintained in RPMI 1640 plus 10% fetal calf serum with daily medium changes. Reverse transcriptase activity of pelleted virions was measured as previously described (33).

**Radioimmunoprecipitation.** For measurement of viral protein expression in infected Jurkat cells,  $5 \times 10^6$  cells were metabolically labeled with 100 μCi (each) of [<sup>35</sup>S]cysteine and [<sup>35</sup>S]methionine per ml for 16 h. Virions were harvested for immunoprecipitation by centrifugation of supernatants for 10 min at  $1,500 \times g$  to remove cell debris and ultracentrifugation at  $12,000 \times g$  for 1 h. Cells or virions were lysed in RIPA lysis buffer, and HIV-1 proteins were immunoprecipitated with AIDS patient serum and analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (23). The human monoclonal antibody (50-69) to gp41 was obtained from the National Institutes of Health AIDS Research and Reference Reagent Program and was a donation of Susan Zolla-Pazner. For measurement of envelope glycoprotein expression in transfected COS-1 cells,  $3 \times 10^6$  cells were metabolically labeled with 100 μCi of [<sup>35</sup>S]methionine per ml for 16 h at 48 h posttransfection with 5 μg each of an envelope expressor plasmid and the *rev* expressor plasmid by the DEAE-dextran method (5). The labeled cells were lysed in 10 mM Tris-HCl (pH 7.4)-150 mM NaCl-5% Triton X-114 at 0°C, and the detergent phase containing integral membrane proteins was isolated by phase separation at 30°C, dissolved in RIPA lysis buffer, clarified, and immunoprecipitated as previously described (3, 11). Virions produced in COS-1 cells were pelleted as described above following labeling of transfected COS-1 cells for 16 h and chasing for 3 h with medium containing excess unlabeled methionine and cysteine.

**PCR assay for determining the efficiency of provirus formation following acute infection.** Fresh virus stocks were prepared from supernatants of infected Jurkat cell cultures by centrifugation at  $1,500 \times g$  for 10 min and filtration (0.45-μm-pore-size filter) to remove cell debris. Virus stocks were treated with DNase (Worthington) (2 μg/ml) for 20 min at room temperature to eliminate plasmid DNA contamination (41). Heat-inactivated virus control supernatants were incubated for 90 min at 60°C. Jurkat cells ( $10^7$ ) were incubated with HXB2 or HXB2Δ(840-856) for 7 h at 37°C. After 7 h,  $2 \times 10^6$  cells were harvested by centrifugation, washed twice, lysed in 50 μl of 0.2% Nonidet P-40, and boiled for 15 min. The cell lysate was clarified by centrifugation at  $12,000 \times g$  for 2 min and stored at -20°C. For polymerase chain reactions (PCR), the DNA in 2.5 μl of the cell lysate (100,000 cell equivalents) was amplified by using the HIV-1 LTR R/U5 primers 5' GGCTAACTAGGGAACCCACTG 3' and 5' CTGCTAGAGATTTTCCACACTGAC 3', which are similar to the AA55 and M667 primer pair previously described (41). PCR reactions were performed according to the manufacturer's instructions (Perkin-Elmer Cetus Corp.) for 33 cycles of 94°C for 1 min, 56°C for 1 min, and 72°C for 1 min. The PCR products were analyzed by electrophoresis on 2% agarose gels.

**Replication complementation assay for measuring the rep-**

**licative potential of mutant envelope glycoproteins.** A transient *trans*-complementation assay was used to assess the replicative potential of mutant envelope glycoproteins in a single round of virus replication (17). Cell-free virus transmission was assessed by cotransfecting COS-1 cells by the DEAE-dextran method (5) with 5 μg of an envelope expressor plasmid and 5 μg of pHXBΔenvCAT. At 48 to 72 h after transfection, the COS-1 cell supernatants were filtered (0.45-μm-pore-size filter) and the reverse transcriptase activity was measured (33). Equivalent reverse transcriptase units of cell-free supernatants were added to  $5 \times 10^6$  Jurkat T lymphocytes. The Jurkat cells were incubated for 48 to 72 h and then assayed for CAT activity. In a similar *trans*-complementation assay, both cell-free and cell-to-cell virus transmission were measured by direct transfection of Jurkat T lymphocytes as previously described (17).

**Syncytium formation assays.** The ability of the mutant glycoproteins to mediate the formation of syncytia was assessed in COS-1 cells by cotransfection of 5 μg of the envelope expressor plasmid and 5 μg of the *rev* expressor plasmid followed by cocultivation with CD4-positive SupT1 lymphocytes for 6 h at 60 h posttransfection (38). The syncytium-forming abilities of the mutant envelope glycoproteins were assessed in Jurkat-*tat* cells by cotransfecting 8 μg of the envelope-expressing plasmid and 8 μg of the *rev* expressor plasmid into Jurkat *tat* cells and scoring syncytia at 60 h posttransfection (11).

**Expression of envelope glycoproteins on the cell surface.** COS-1 cells cotransfected with an envelope expressor plasmid and the *rev* expressor plasmid were metabolically labeled with [<sup>35</sup>S]methionine for 16 h at 48 h posttransfection. The intact labeled cells were washed twice with ice-cold phosphate-buffered saline (PBS) containing 2% heat-inactivated fetal calf serum, incubated with a 1:100 dilution of AIDS patient serum reactive with the envelope glycoproteins at 4°C for 30 min, rinsed twice with PBS containing 2% fetal calf serum, and lysed in RIPA lysis buffer containing 2.5 μg of unlabeled gp120 (American BioTechnologies, Inc.) at 4°C. The cell lysates were clarified by ultracentrifugation, and bound envelope glycoproteins were immunoprecipitated by incubation with protein A-Sepharose as described above.

**Soluble-CD4 inhibition of syncytium formation and virus replication.** The effect of soluble CD4 on virus entry was measured by producing the recombinant HXBΔenvCAT proviruses in COS-1 cells as described above and preincubating equivalent amounts of virus as determined by measuring reverse transcriptase activity with different concentrations of full-length soluble CD4 (American BioTechnologies) for 1 h at 37°C prior to infection of Jurkat lymphocytes (38).

## RESULTS

**Effects of a deletion at the carboxyl terminus of gp41 on HIV-1 replication in Jurkat lymphocytes.** A mutation which deletes residues 840 to 856 (28) at the carboxyl terminus of gp41 was introduced into an infectious HIV-1 provirus on plasmid pHXB2. To compare the replication rate of this mutant virus with that of the wild-type virus, the pHXB2 and pHXB2Δ(840-856) plasmids were transfected into Jurkat cells and virus replication was monitored by measuring reverse transcriptase activity in the culture supernatants. The Jurkat cultures transfected with the pHXB2Δ(840-856) plasmid demonstrated slowed virus replication, with a 5-day lag in the time required to reach peak reverse transcriptase activity compared with that of the wild-type virus (Fig. 1), which is consistent with previous studies (9, 26, 37). When



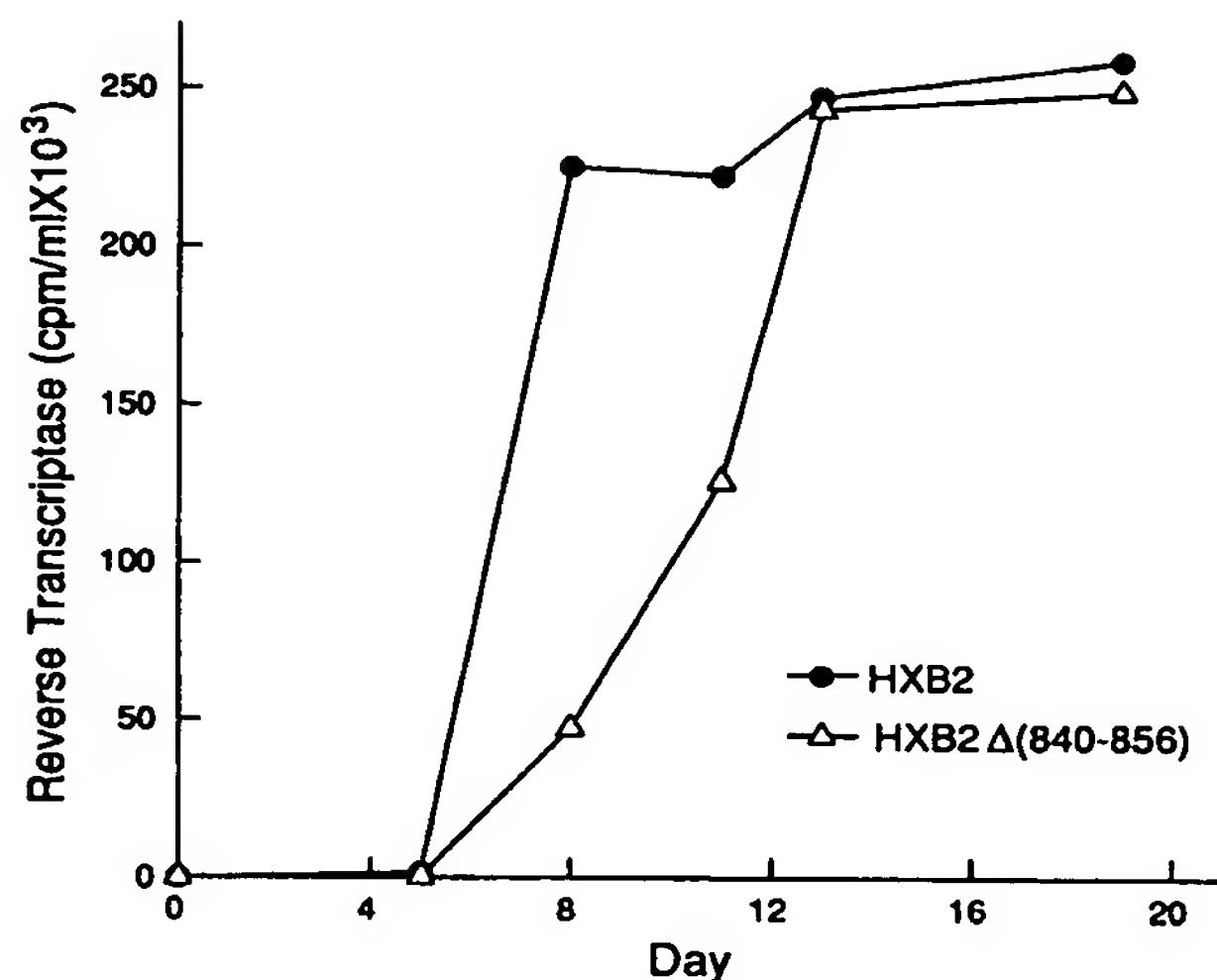


FIG. 1. Replication of HXB2 and HXB2Δ(840-856) viruses in Jurkat cells. Curves show reverse transcriptase activity in supernatants of Jurkat cell cultures transfected with 10 μg of pHXB2 or pHXB2Δ(840-856) DNA. The results are typical of those seen in at least three independent experiments.

similar levels of virus production were achieved, the number of syncytia formed in the pHXB2- and pHXB2Δ(840-856)-transfected cultures and the viability of these cultures were indistinguishable (data not shown).

To examine the effects of a deletion at the carboxyl terminus of gp41 on *gag* and *env* protein synthesis, processing, secretion, and incorporation into virions in Jurkat cells, steady-state and pulse-chase labeling experiments were performed at a time when virus production in HXB2- and HXB2Δ(840-856)-infected cultures was similar. At 14 to 21 days after transfection, equal numbers of viable cells were metabolically labeled with [<sup>35</sup>S]cysteine and [<sup>35</sup>S]methionine for 16 h and viral proteins were immunoprecipitated from cell lysates, supernatants, and pelleted virions. The steady-

state levels of viral *gag* and *env* proteins were similar in the pHXB2- and pHXB2Δ(840-856)-transfected cell lysates, supernatants, and pelleted virions (Fig. 2A, B, and C and data not shown), although slightly less cell-associated gp160 was observed in the pHXB2Δ(840-856)-transfected culture (Fig. 2A). Pulse-chase analysis of infected Jurkat cell cultures demonstrated similar synthesis, processing, and release of envelope glycoproteins (Fig. 2D and E). However, the level of the mutant gp160 in cell lysates was approximately 50% of the wild-type level after a 2-h or longer chase (Fig. 2D). This decrease in the level of gp160 at the longer chase times was not accompanied by an increase in the level of cell-associated gp120 glycoprotein. The gp120 glycoprotein was released into the cell supernatants after a 1-h chase in both the wild-type and mutant cultures (Fig. 2E). These results suggest that deletion of residues 840 to 856 at the carboxyl terminus of gp41 results in decreased stability of the gp160 glycoprotein in infected Jurkat cells. Nonetheless, the synthesis, processing, secretion, and incorporation into virions of the mutant HXB2Δ(840-856) envelope glycoproteins appear to be comparable to those of the wild-type glycoproteins.

**Deletion of the carboxyl terminus of gp41 reduces the efficiency of virus entry.** To determine whether the cytoplasmic domain of gp41 has an effect on the efficiency of virus entry, HXB2 or HXB2Δ(840-856) virus stocks were used to infect Jurkat cells and the efficiency of HIV-1 provirus formation at 7 h after acute infection was measured by PCR amplification of the proviral DNA. Uninfected Jurkat cells and Jurkat cells incubated with heat-inactivated virus controls were used as controls (41). The HIV-1 LTR R/U5 primer pair used for PCR amplification detects the first region of the viral DNA made during reverse transcription and should detect virtually all HIV-1 DNA made, including partial DNA transcripts (41). Mapping of restriction sites within the amplified DNA product with *HindIII*, *HinfII*, *XhoII*, and *MboI* restriction enzymes, which cut at internal sites, confirmed that the 140-bp fragment amplified by these primers was the correct HIV-1 LTR fragment (data not shown). Following acute infection of Jurkat cells, formation of the HXB2Δ(840-856) provirus as detected by PCR amplification of viral DNA was significantly reduced compared

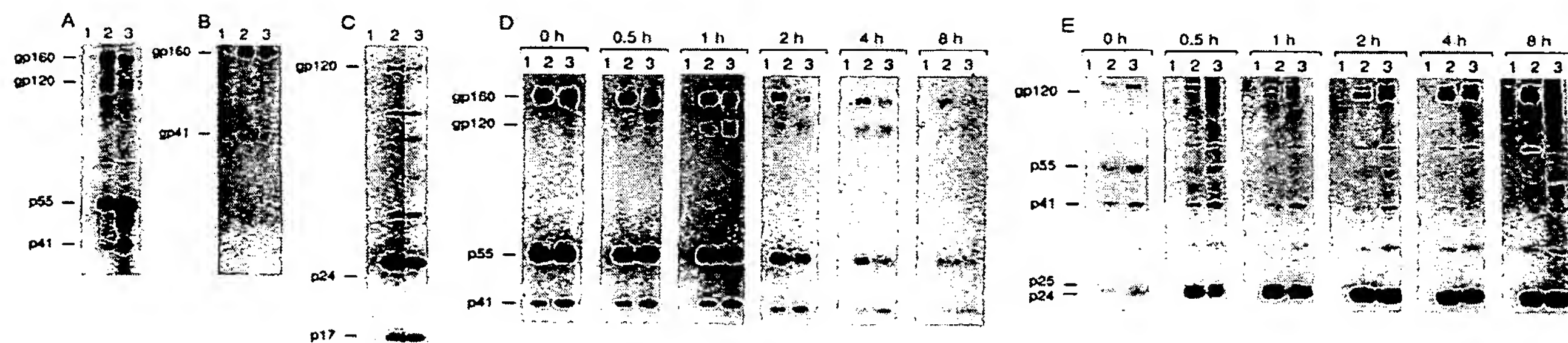


FIG. 2. HIV-1 *env* and *gag* proteins in HXB2- and HXB2Δ(840-856)-infected Jurkat cells and incorporated into virions. Immunoprecipitated cell lysates (A, B, D), virions (C), and supernatants (E) were assayed at days 14 to 21 following transfection of Jurkat cells with no DNA (lanes 1), 10 μg of pHXB2 (lanes 2), or 10 μg of pHXB2Δ(840-856) (lanes 3). Cells were metabolically labeled with [<sup>35</sup>S]cysteine and [<sup>35</sup>S]methionine for 16 h (A, B, C) or by a pulse-chase protocol (D, E). The gp160, gp120, and gp41 envelope glycoproteins and the p55, p24, and p17 *gag* products are marked. (A) Immunoprecipitated cell lysates with AIDS patient serum analyzed by 4 to 12% gradient SDS-PAGE; (B) immunoprecipitated cell lysates with a human monoclonal antibody (50-69) that recognizes gp41 analyzed by 12.5% SDS-PAGE; (C) immunoprecipitated pelleted virions with AIDS patient serum analyzed by 12.5% SDS-PAGE; (D and E) pulse-chase analysis of cell lysates (D) and supernatants (E) immunoprecipitated with AIDS patient serum and analyzed by 4 to 12% gradient SDS-PAGE. Cells were labeled for 10 min and chased for the indicated times (in hours) with medium containing excess unlabeled cysteine and methionine.



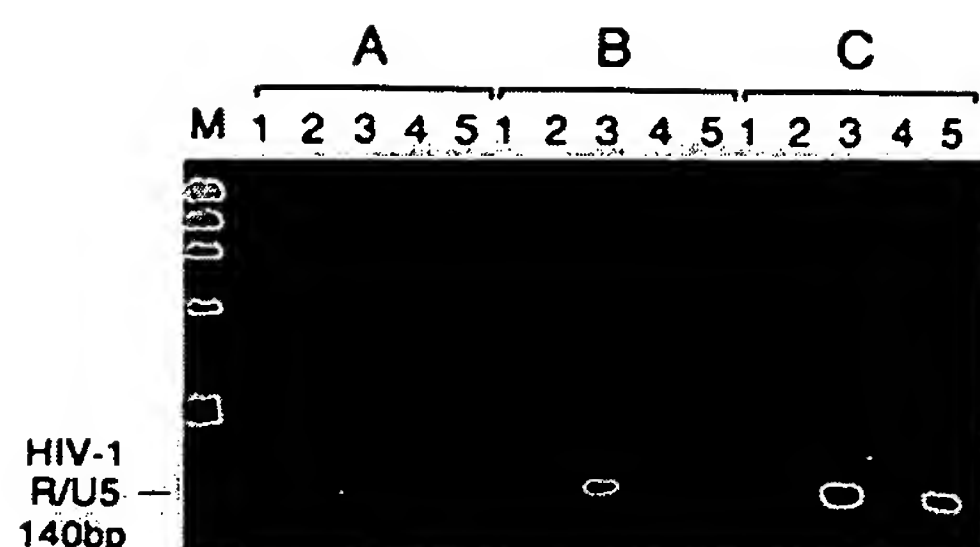


FIG. 3. Detection of HIV-1 provirus formation in Jurkat cells by PCR amplification of viral DNA following acute infection with HXB2 or HXB2Δ(840-856) viruses. PCR amplification was performed by using the HIV-1 LTR R/U5 primer pair M667-AA55 (41) for 33 cycles to amplify DNA from 100,000 Jurkat cell equivalents harvested 7 h after acute infection with 4,000 (A), 20,000 (B), or 100,000 (C) cpm of reverse transcriptase units of no virus (lane 1), heat-inactivated HXB2 (lane 2), HXB2 (lane 3), heat-inactivated HXB2Δ(840-856) (lane 4), or HXB2Δ(840-856) (lane 5) per ml. The 140-bp PCR product was analyzed by electrophoresis on 2% agarose gels and visualized with ethidium bromide. A DNA standard (*PhiX-HaeIII* digest) is shown in lane M.

with that of wild type (Fig. 3). As shown in Fig. 3, the effect of the mutation on the efficiency of virus entry was most marked at the lower multiplicities of infection (4,000 and 20,000 cpm/ml). These results suggest that the carboxyl terminus of gp41 has an effect on the efficiency of virus entry during an early stage in the virus life cycle prior to formation of the DNA provirus.

**Mutational analysis of sequences in the cytoplasmic domain of gp41 required for replication.** The cytoplasmic domain of gp41 contains a strongly hydrophilic region (amino acids 724 to 745) followed by a region of alternating hydrophobicity and hydrophilicity (amino acids 746 to 856) (4, 25). The latter region also contains two segments with the potential to form amphipathic  $\alpha$  helices (amino acids 770 to 794 and 824 to 856) (1, 39). A series of mutants containing deletions in the cytoplasmic domain was constructed in the pSVIIIenv envelope expressor plasmid (Fig. 4). Because many of these mutants overlap the *tat* or *rev* second open reading frames, *rev* or *rev* plus *tat* was supplied in *trans* by cotransfecting either the *rev* expressor plasmid or pHXB2ΔenvCAT, respectively. Of note, the truncation in the Δ(726-856) mutant is similar to the truncation found in some isolates of SIV and HIV-2 (4, 10). The effect of the Δ(814-856) mutation in the

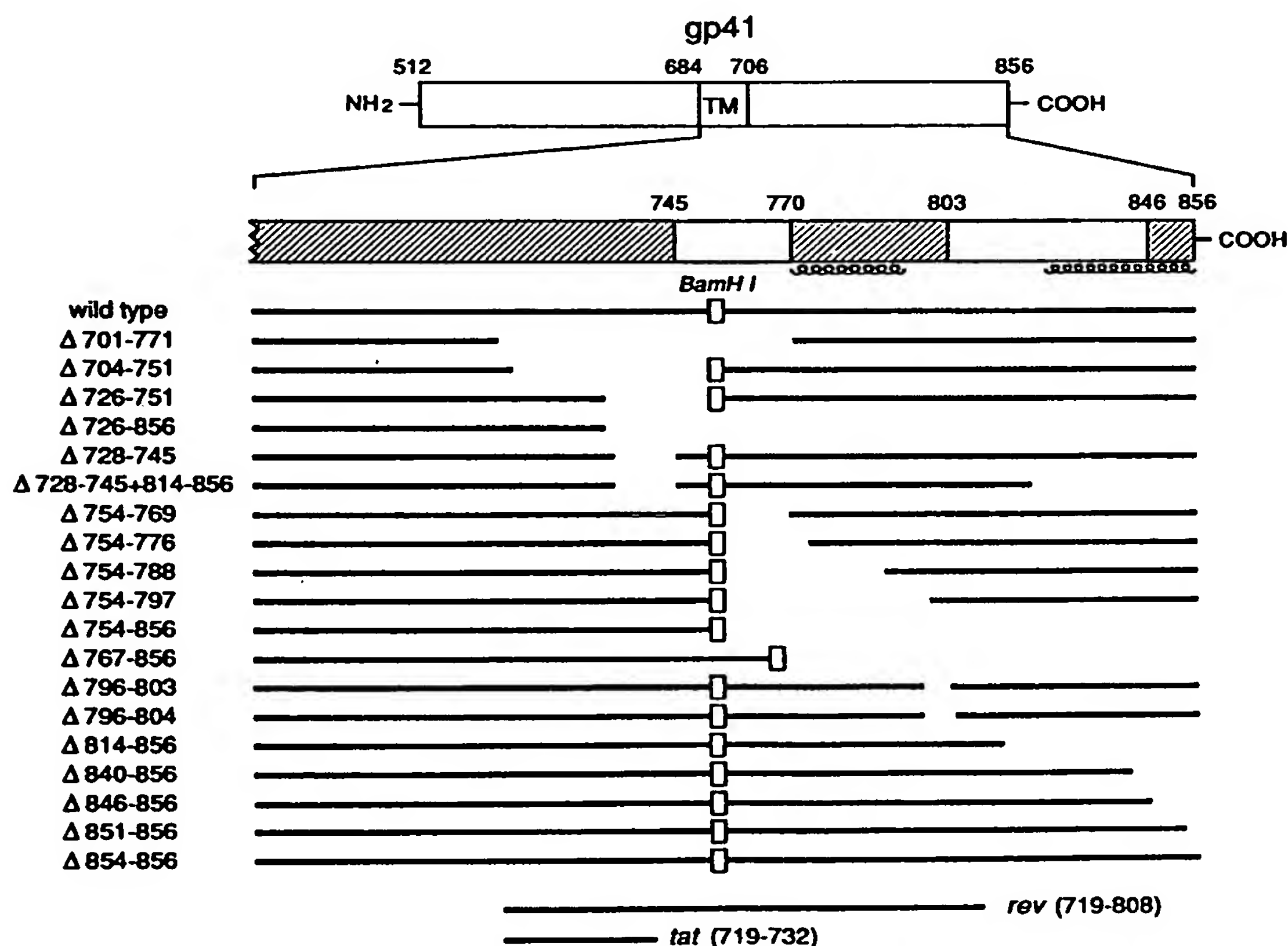


FIG. 4. Diagram of the mutant HIV-1 gp41 transmembrane glycoproteins. The wild-type HIV-1 gp41 transmembrane glycoprotein is depicted at the top, with the hydrophobic transmembrane region designated TM. The cross-hatched regions are highly conserved among HIV-1 isolates (28). Regions of potential amphipathic  $\alpha$  helices at residues 770 to 794 and 824 to 856 (39) are indicated by coils. The positions of the second open reading frames of *tat* and *rev* are shown.

TABLE 1. Phenotypes of gp41 cytoplasmic domain mutants<sup>a</sup>

Mutant(s) <sup>b</sup>	Replication complementation <sup>c</sup>		Relative cell-associated gp120 <sup>c</sup>	Processing index <sup>d</sup>	Syncytium formation <sup>e</sup>	
	COS-1	Jurkat			COS-1	Jurkat
None	2	5	0	ND	9	5
Wild type	100	100	100	1.0	100	100
Δ701-771	8	25	27	0.6	38	10
Δ704-751	38	50	28	0.7	68	96
Δ726-751	32	86	58	0.8	67	60
Δ726-856	28	7	32	1.0	71	83
Δ728-745	31	88	63	0.9	64	56
Δ728-745 and 814-856	9	10	ND	ND	ND	ND
Δ754-769	60	95	ND	ND	ND	ND
Δ754-776	63	37	ND	ND	ND	ND
Δ754-788	69	38	49	0.9	37	63
Δ754-797	31	10	29	0.9	20	33
Δ754-856	6	9	7	1.0	71	106
Δ767-856	6	12	32	1.5	30	92
Δ796-803	50	14	56	1.1	72	ND
Δ796-804	4	7	33	1.2	65	ND
Δ814-856	24	17	57	1.1	47	126
Δ840-856	34	24	48	1.1	63	135
Δ846-856	47	24	96	1.0	102	92
Δ851-856	50	28	121	0.9	80	122
Δ854-856	89	32	111	1.0	80	86

<sup>a</sup> Data reported represent the means of at least two or three independent experiments. ND, not done.<sup>b</sup> The number of the mutant refers to the amino acid residues of the HXB2 strain of HIV-1, where 1 is the initial methionine (31).<sup>c</sup> Percentage of the wild-type value. Values for replication complementation represent the CAT conversion in target Jurkat cells for the mutant envelope glycoprotein relative to the value for the wild type. Values shown for cell-associated gp120 were determined by gel densitometry of radioimmunoprecipitates analyzed by SDS-PAGE and autoradiography.<sup>d</sup> A measure of the conversion of mutant gp160 to gp120 relative to that of the wild-type glycoprotein. The amounts of gp160 and gp120 were determined by gel densitometry of autoradiograms of SDS-PAGE gels. The processing index was calculated by the following formula: processing index = [(total gp120)<sub>mutant</sub> × (gp160)<sub>wild type</sub>] / [(gp160)<sub>mutant</sub> × (total gp120)<sub>wild type</sub>].

context of an infectious HIV-1 provirus has already been described elsewhere (37).

To identify sequences in the cytoplasmic domain that are required for efficient viral replication, the ability of the mutant envelope glycoproteins to support cell-free virus transmission was determined by using a transient complementation assay (17). In this assay, COS-1 cells cotransfected with an envelope expressor plasmid and an envelope-defective provirus expressing the CAT gene produce recombinant virions that are used to infect Jurkat T lymphocytes (17). The efficiency of cell-free transmission of the recombinant viruses is determined by measuring CAT activity in the infected Jurkat culture. Most deletions within the carboxy-terminal 160 amino acids of gp41 reduced replication complementation of cell-free virus transmission to 4 to 69% of the wild-type value (Table 1). To determine whether similar regions are required for replication in Jurkat cells under conditions in which most of the virus transmission occurs by cell-to-cell spread, the abilities of the mutant glycoproteins to support a single round of virus replication were examined by using a similar transient complementation assay in transfected Jurkat cell cultures (17). The hydrophilic region extending from amino acids 726 to 751 and the immediately carboxyl region (amino acids 754 to 769) were relatively dispensable for replication in Jurkat cells (Table 1). Otherwise, the regions in the cytoplasmic domain required to support the transmission of viruses produced in COS-1 cells and Jurkat cells were similar. These results suggest that the function of some regions of the cytoplasmic domain depends on the cell type in which the envelope

glycoproteins are expressed or on the mode of virus transmission.

Deletions in the cytoplasmic domain of gp41 are associated with decreased envelope glycoprotein stability in COS-1 cells. To determine whether mutations in the cytoplasmic domain of gp41 alter envelope glycoprotein synthesis or processing in COS-1 cells, mutant envelope glycoproteins expressed in COS-1 cells were immunoprecipitated with AIDS patient serum and analyzed on SDS-polyacrylamide gels. The steady-state levels of mutant glycoproteins with deletions within residues 846 to 856 at the carboxyl terminus were similar to those of the wild-type glycoproteins (Fig. 5A and Table 1). The steady-state levels of mutant glycoproteins containing deletions amino terminal to residue 846 were reduced (Fig. 5A and Table 1). Two mutants containing deletions overlapping the transmembrane domain (amino acids 701 to 777 and 704 to 751) exhibited a processing defect (Fig. 5A and Table 1). The levels of wild-type and mutant gp120 in the transfected COS-1 cell supernatants correlated with the levels of cell-associated gp120 (data not shown). When labeling COS-1 cells for 16 h was followed by a 3-h chase with excess cold methionine and cysteine, the levels of mutant envelope glycoproteins that were low in the steady state were further reduced (Fig. 5B and Table 2). This result suggests that deletion of residues amino terminal to amino acid 846 results in decreased envelope glycoprotein stability in COS-1 cells.

The cytoplasmic domain of gp41 is not required for incorporation of envelope glycoproteins into virions. To determine whether the cytoplasmic domain of gp41 is required for the

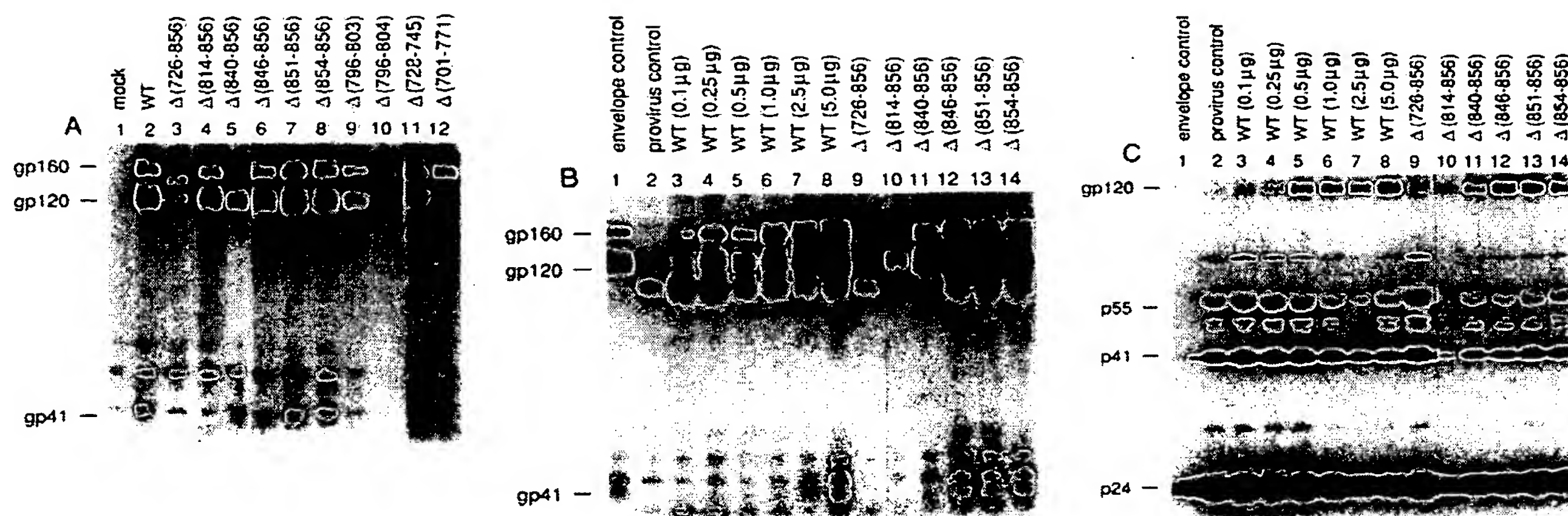


FIG. 5. Wild-type and mutant HIV-1 envelope glycoproteins in COS-1 cell lysates (A, B) and virion pellets (C). Cell lysates and virions were immunoprecipitated with AIDS patient serum and analyzed by 10% (A, B) or 12.5% (C) SDS-PAGE. (A) Steady-state levels of mutant envelope glycoproteins expressed in COS-1 cells transfected with 5  $\mu$ g of the *rev* expressor plasmid plus no DNA (lane 1), 5  $\mu$ g of wild-type envelope expressor plasmid (lane 2), or 5  $\mu$ g of the mutant envelope expressor plasmids (lanes 3 to 12). (B) Expression of wild-type (lanes 1 and 3 to 8) and mutant (lanes 9 to 14) envelope glycoproteins in COS-1 cells after 16 h of labeling and 3 h of chase. COS-1 cells were transfected with 5  $\mu$ g of the wild-type envelope expressor plasmid (lane 1), pHXB $\Delta$ env $\Delta$ EcoCAT (lane 2), or pHXB $\Delta$ env $\Delta$ EcoCAT plus 0.1, 0.25, 0.5, 1.0, 2.5, or 5.0  $\mu$ g of the wild-type envelope expressor plasmid (lanes 3 to 8, respectively) or 5  $\mu$ g of the mutant envelope expressor plasmids (lanes 9 to 14). (C) Virion-associated envelope glycoproteins. The lanes are the same as in panel B. The positions of the gp120 envelope glycoprotein and p55, p41, and p24 *gag* products are marked.

TABLE 2. Effect of mutations in the gp41 cytoplasmic domain on virion-associated envelope glycoproteins<sup>a</sup>

Mutant <sup>b</sup>	Relative cell-associated gp120 after 3-h chase <sup>c</sup>	Virion-associated gp120 (gp120/p24) <sup>c</sup>	Virion association index <sup>d</sup>
Wild type (0.1 $\mu$ g) <sup>a</sup>	9	15	1.7
Wild type (0.25 $\mu$ g)	16	22	1.4
Wild type (0.5 $\mu$ g)	22	46	2.0
Wild type (1.0 $\mu$ g)	39	35	0.9
Wild type (2.5 $\mu$ g)	56	57	1.0
Wild type (5.0 $\mu$ g)	100	100	1.0
$\Delta$ 701-771	4	25	5.6
$\Delta$ 704-751	18	25	2.6
$\Delta$ 726-856	4	15	4.0
$\Delta$ 728-745	58	70	1.4
$\Delta$ 754-788	12	20	1.7
$\Delta$ 754-856	4	4	ND
$\Delta$ 767-856	15	32	2.6
$\Delta$ 814-856	39	33	1.0
$\Delta$ 840-856	62	42	0.8
$\Delta$ 846-856	100	76	0.8
$\Delta$ 851-856	84	72	0.9
$\Delta$ 854-856	104	83	0.8

<sup>a</sup> Data reported represent the means of at least two independent experiments.

<sup>b</sup> The numbers in parentheses refer to the amount of wild-type envelope expressor plasmid DNA used for transfection. For all of the mutant glycoproteins, 5  $\mu$ g of plasmid DNA was used.

<sup>c</sup> Percentage of the wild-type value at 5  $\mu$ g of wild-type transfected plasmid DNA. The values shown were determined by gel densitometry of radioimmunoprecipitates analyzed by SDS-PAGE and autoradiography.

<sup>d</sup> Calculated to normalize the level of virion-associated gp120 for the level of cell-associated gp120. The virion-association index was calculated by using the formula [(virion-associated gp120)<sub>mutant</sub>  $\times$  (p24)<sub>wild type</sub>  $\times$  (cell-associated gp120)<sub>wild type</sub>] / [(p24)<sub>mutant</sub>  $\times$  (virion-associated gp120)<sub>wild type</sub>  $\times$  (cell-associated gp120)<sub>mutant</sub>]. ND, not done.

efficient incorporation of HIV-1 envelope glycoproteins into virions, the effect of deletions on the level of virion-associated envelope glycoproteins was examined. COS-1 cells cotransfected with an envelope expressor plasmid and pHXB $\Delta$ env $\Delta$ EcoCAT were used to produce recombinant virions. Pelleted virions were immunoprecipitated with AIDS patient serum and analyzed on SDS-polyacrylamide gels. The amount of transfected wild-type envelope expressor plasmid DNA was incrementally increased from 0.1 to 5.0  $\mu$ g to allow comparison of the level of wild-type and mutant virion-associated gp120 under conditions in which the levels of cell-associated envelope glycoproteins were similar (Table 2 and Fig. 5B and C). A virion association index was calculated to normalize the level of virion-associated gp120 for the level of cell-associated gp120 (Table 2). Deletions amino terminal to residue 846 caused reductions in the level of virion-associated gp120 to 70% or less of the wild-type level (Fig. 5C and Table 2). However, the virion association index of the mutant glycoproteins was similar to or greater than that of the wild-type envelope glycoproteins under conditions in which the levels of cell-associated gp120 were similar (Table 2). Moreover, the relative incorporation of gp120 into virions was generally more efficient when the levels of wild-type or mutant cell-associated gp120 were relatively low (Table 2). These results demonstrate that the cytoplasmic domain of gp41 is not required for the efficient incorporation of the gp120 envelope glycoprotein into virions. It can also be inferred that the incorporation of the gp41 glycoprotein is not defective, since the cell or virus association of the gp120 glycoprotein depends on an interaction with gp41 (24).

To determine whether the observed decrease in the absolute amount of virion-associated mutant envelope glycoprotein was the sole cause of the replication defect in virions made in COS-1 cells, replication complementation by the wild-type and mutant envelope glycoproteins was compared under conditions in which levels of virion-associated gp120



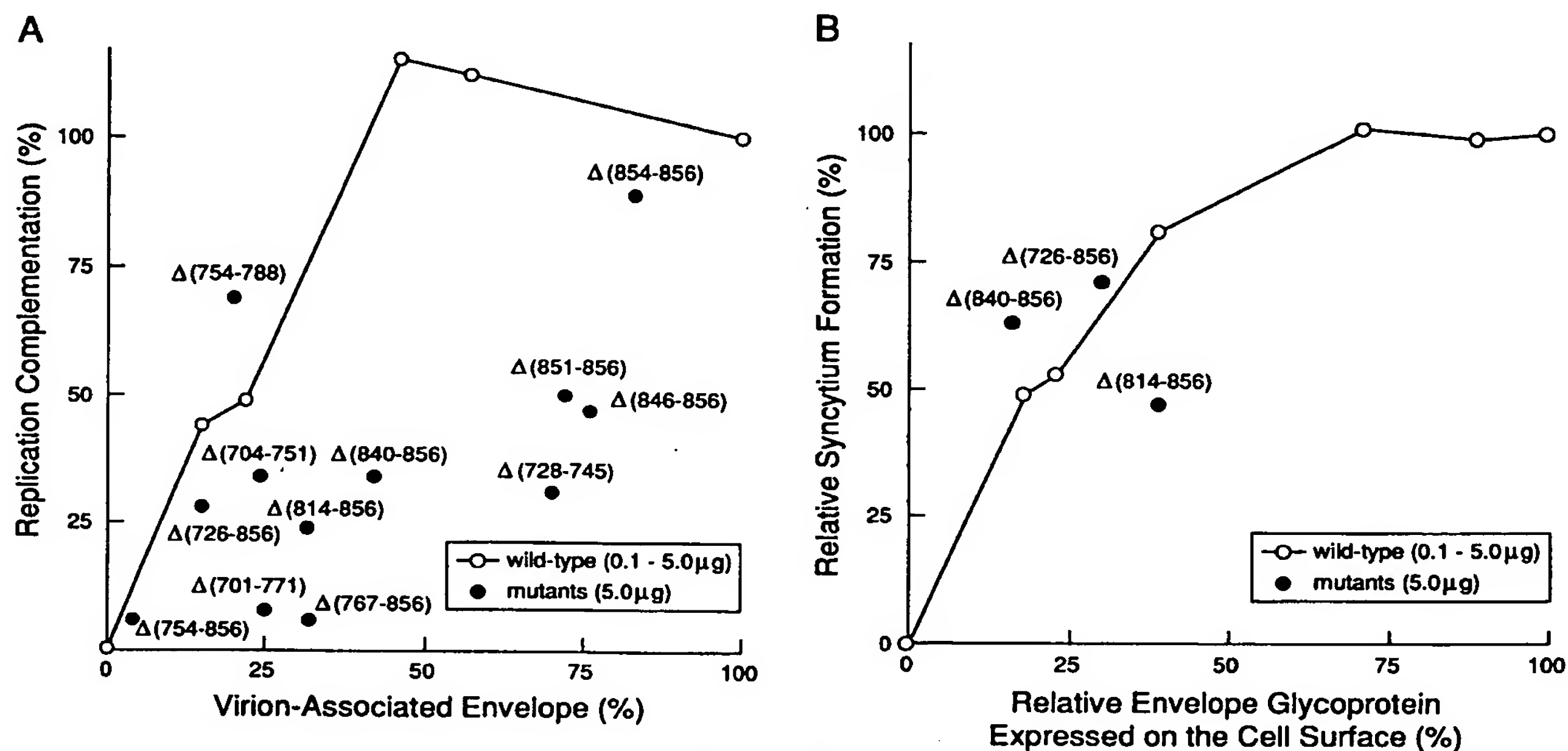


FIG. 6. (A) Levels of virion-associated envelope glycoproteins and replication complementation. Values for virion-associated envelope glycoproteins were derived from Table 2. Values for replication complementation were determined by using the replication complementation assay in COS-1 cells (Table 1). All values are expressed as a percentage of the value for the wild-type envelope glycoprotein at 5  $\mu$ g of transfected plasmid DNA. Open circles represent the values, in consecutive order, for the wild-type envelope glycoproteins when 0.1, 0.25, 0.5, 2.5, or 5.0  $\mu$ g of envelope expressor plasmid DNA was transfected. For the mutant envelope glycoproteins, 5  $\mu$ g of envelope expressor plasmid DNA was transfected. (B) Levels of envelope glycoproteins expressed on the cell surface and syncytium-forming ability. The expression of envelope glycoproteins on the cell surface was determined by immunoprecipitation of cell surface-accessible envelope glycoproteins following incubation of intact COS-1 cells with AIDS patient serum, analysis by SDS-PAGE, and quantitation by gel densitometry. All values are expressed as a percentage of the value for the wild-type envelope glycoproteins at 5  $\mu$ g of transfected plasmid DNA. Open circles represent values, in consecutive order, for the wild-type envelope glycoproteins when 0.1, 0.25, 0.5, 1.0, 2.0, or 5.0  $\mu$ g of plasmid DNA was transfected. For the mutant envelope glycoproteins, 5.0  $\mu$ g of plasmid DNA was transfected.

were similar. The relationship between the levels of virion-associated wild-type glycoprotein and degree of replication complementation was nonlinear, with saturation of the latter value when the level of virion-associated glycoprotein was only 35% of the maximum value (Fig. 6A). As shown in Fig. 6A, relatively low levels of wild-type virion-associated envelope glycoprotein can mediate virus entry efficiently. In contrast, most mutant envelope glycoproteins containing deletions within the cytoplasmic domain exhibited a replication defect that could not be solely attributed to the reduced level of virion-associated envelope glycoproteins (Fig. 6A).

**The gp41 cytoplasmic domain is not required for syncytium formation.** To determine whether a defect in fusion was the cause of the reduced replication complementation ability of the mutant envelope glycoproteins produced in COS-1 cells, COS-1 cells were cotransfected with a wild-type or mutant envelope expressor plasmid and cocultured with CD4-positive SupT1 cells, and syncytium-forming ability was scored. The syncytium-forming abilities of most of the mutant glycoproteins expressed in COS-1 cells were less than those of the wild-type glycoproteins (Table 1). Since previous studies had demonstrated that syncytium formation in other cell types did not require the cytoplasmic domain (7, 24), the syncytium-forming abilities of the mutant envelope glycoproteins were also examined in Jurkat-*tat* cells (11, 18). In Jurkat-*tat* cells, the syncytium-forming abilities of the mutant glycoproteins were usually equal to or greater than

those of the wild-type glycoproteins (Table 1). These results suggest either that the syncytium-forming ability of the mutant glycoproteins is dependent on the cell type in which the envelope glycoproteins are expressed or that syncytium formation in COS-1 cells might be reduced as a consequence of decreased expression of the mutant envelope glycoproteins on the cell surface.

To determine whether the decrease in syncytium formation observed in COS-1 cells could be solely attributed to reduced expression of envelope glycoproteins on the cell surface, the expression of selected mutant envelope glycoproteins on the cell surface was determined by immunoprecipitation, analysis on SDS-polyacrylamide gels, and quantitation by gel densitometry. The syncytium-forming abilities of the wild-type and mutant envelope glycoproteins were compared under conditions in which similar levels were expressed on the cell surface. For the  $\Delta(726-856)$ ,  $\Delta(814-856)$ , and  $\Delta(840-856)$  mutant glycoproteins, the level of envelope glycoproteins expressed on the cell surface was not significantly different from the wild-type level when the levels of intracellular envelope glycoproteins were similar (Table 2 and Fig. 6B). Furthermore, the syncytium-forming abilities of the mutant envelope glycoproteins made in COS-1 cells were similar to that of wild type when similar levels of envelope glycoproteins were expressed on the cell surface (Fig. 6B). These results demonstrate that the cytoplasmic domain is not required for the cell surface expres-

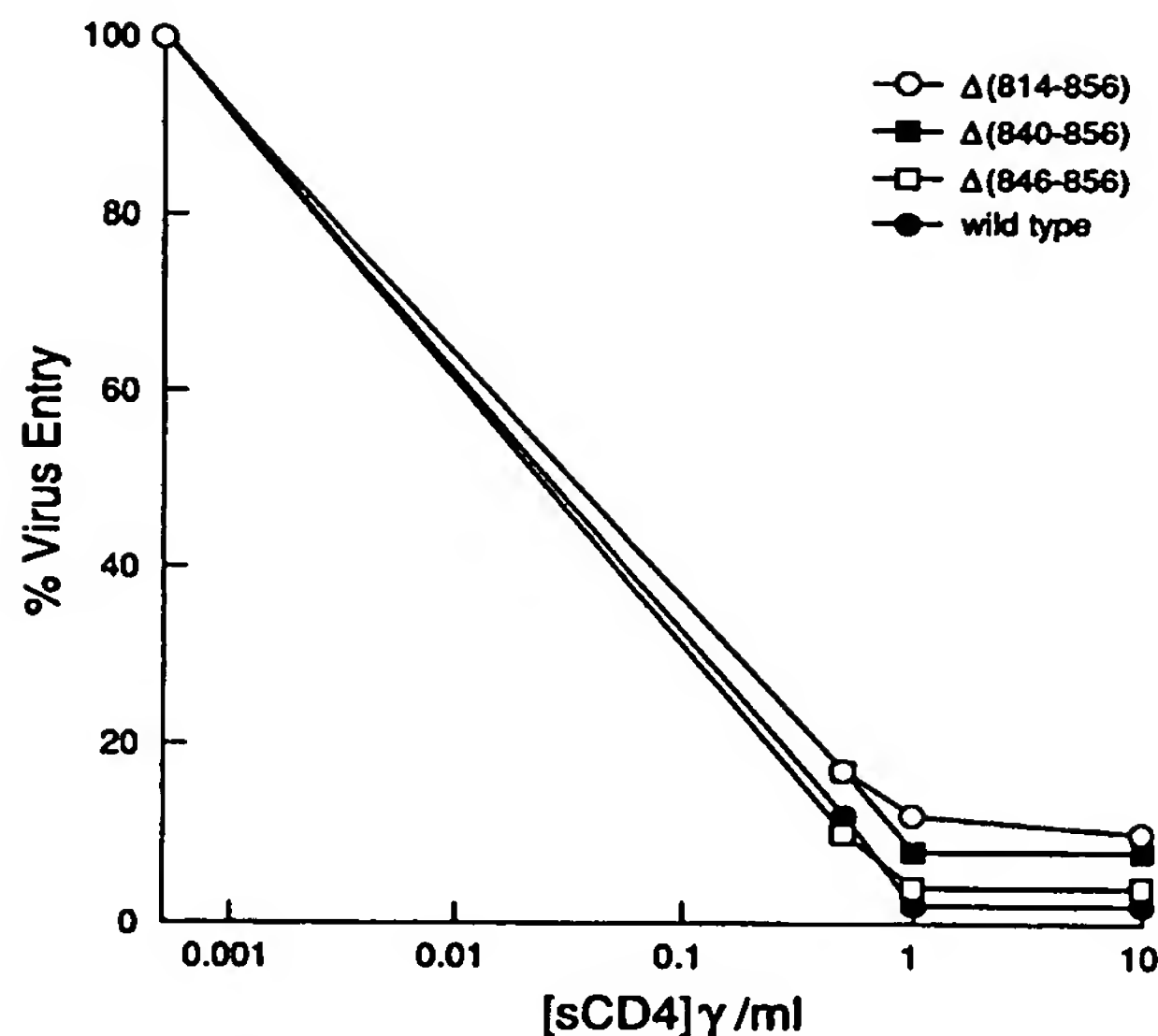


FIG. 7. Effect of mutations in the cytoplasmic domain on the sensitivity of replication complementation to soluble-CD4 (sCD4) inhibition. The effects of different soluble-CD4 concentrations on the CAT activity transferred to Jurkat lymphocytes by recombinant virions carrying the wild-type or mutant envelope glycoproteins are shown. Each value represents the percentage of CAT activity observed for a given mutant in the presence of soluble CD4 relative to the activity observed for the mutant in the absence of soluble CD4.  $\gamma$ , micrograms.

sion or the fusion function of envelope glycoproteins expressed in COS-1 cells.

**Soluble-CD4 sensitivity of the mutant envelope glycoproteins.** HIV-1 envelope glycoproteins that exhibit more than a twofold decrease in relative CD4-binding ability are less sensitive than the wild-type glycoprotein to soluble CD4 inhibition of virus entry (38). To investigate the possibility that deletions in the cytoplasmic domain of gp41 affect the CD4-binding ability of the functional envelope glycoprotein multimer, the effect of mutations in the cytoplasmic domain on soluble CD4 inhibition of virus entry was examined. Equivalent amounts of recombinant virions produced in COS-1 cells as determined by measuring reverse transcriptase activity were incubated with different concentrations of soluble CD4 for 1 h at 37°C prior to infection of Jurkat target cells. The  $\Delta(814-856)$ ,  $\Delta(840-856)$ , and  $\Delta(846-856)$  mutants exhibited the same sensitivity to soluble-CD4 inhibition of virus entry as the wild-type envelope glycoprotein (Fig. 7). From these results, it can be inferred that deletions in the cytoplasmic domain of gp41 do not significantly alter the CD4-binding ability of the native, multimeric envelope glycoproteins. This conclusion is consistent with previous observations that alterations in the gp41 cytoplasmic domain do not affect the CD4-binding ability of the monomeric soluble form of the gp120 glycoprotein (2, 7, 8, 24).

## DISCUSSION

This study demonstrates that the cytoplasmic domain of the HIV-1 transmembrane glycoprotein has at least two effects on viral replication. A highly conserved region within

residues 840 to 856 at the carboxyl terminus of gp41 is important for efficient virus entry during an early step in the virus life cycle between CD4 binding and formation of the DNA provirus. The entire cytoplasmic domain amino terminal to residue 846 affects the stability of envelope glycoproteins made in COS-1 cells. The entire cytoplasmic domain appears to affect the efficiency of virus entry during the cell-free transmission of viruses made in COS-1 cells, whereas residues 726 to 769 are dispensable for cell-to-cell transmission in Jurkat cells. Whether these differences are a consequence of the cell type in which the envelope glycoproteins are made or reflect differences in the requirement for this domain during cell-free versus cell-to-cell transmission remains to be determined.

The reduced efficiency of HIV-1 provirus formation following acute infection with HXB2 $\Delta(840-856)$  demonstrates that the carboxyl terminus of the HIV-1 transmembrane glycoprotein has an effect on the efficiency of virus entry during an early stage in the virus life cycle prior to formation of the provirus. The steps in the virus life cycle required for virus entry include CD4 binding, fusion of the virus and host cell membranes, and penetration and uncoating of the viral core (13). Deletion of the carboxyl terminus of gp41 is unlikely to affect CD4 binding, since the CD4-binding domains are all contained within the gp120 subunit (24, 27). The lack of requirement for the cytoplasmic domain for syncytium formation and the lack of altered sensitivity to soluble CD4 provide further evidence that deletions in this region do not significantly affect the CD4-binding affinity of the multimeric glycoprotein complex. Deletions in the cytoplasmic domain of gp41 are also not likely to affect the fusion function, since this domain is not required for syncytium formation, as demonstrated here and in previous studies (7, 24). The possibility that the membrane fusion function of only the virion-associated envelope glycoproteins is defective has not been excluded, but reliable methods for measuring the fusion function of the small fraction of virions relevant to infection are not available. The competence of the mutant envelope glycoproteins for CD4 binding and membrane fusion in conjunction with inefficient formation of the DNA provirus raises the possibility that these truncated envelope glycoproteins are defective for a step involved in the uncoating or penetration of the viral core. The location of the gp41 carboxyl terminus within the virion close to the viral core is consistent with the possibility that this region plays a role in virus uncoating or penetration.

Independent of a role in virus replication, the integrity of the gp41 cytoplasmic domain amino terminal to residue 846 is important for the stability of envelope glycoproteins expressed in COS-1 cells. The observation that the stability of gp160 in HXB2 $\Delta(840-856)$ -infected Jurkat cell cultures is decreased suggests that the cytoplasmic domain also has an effect on the stability of the envelope glycoproteins in other cell types. The effect of the cytoplasmic domain on envelope glycoprotein stability appears to be cell type dependent, judging by the observation that the effect is relatively minor in Jurkat cells and the finding that some mutant glycoproteins that showed decreased stability in COS-1 cells were fully functional, and thus presumably stable, in Jurkat cells. These results are consistent with previous studies demonstrating that the stability and processing of the HIV-1 envelope glycoproteins depend on the cell type in which the envelope glycoproteins are expressed (8, 18). In COS-1 cells, approximately 75% of gp160 is cleaved to produce the gp120 subunit, whereas in Jurkat cells, only a small fraction (10 to 20%) of gp160 is cleaved. This dependence of the efficiency

of gp160 cleavage on cell type might, at least in part, account for some of the cell type-specific effects observed in this study, since a decrease in the steady-state level of gp160 would be less deleterious in cells which process gp160 inefficiently. In a previous study, Willey et al. demonstrated that intracellular sorting results in the transport of most uncleaved gp160 to lysosomes, where it is degraded (40). The mechanism by which the cytoplasmic domain affects envelope glycoprotein stability is unknown. One possibility is that deletions in the cytoplasmic domain result in altered transport of gp160, as observed in one study (16), which might affect localization to subcellular compartments where degradation occurs.

The cytoplasmic domain of the transmembrane protein does not appear to be required for the efficient incorporation of the HIV-1 envelope glycoproteins into virions. These findings are consistent with those of previous studies demonstrating that SIV and equine infectious anemia virus containing truncated envelope glycoproteins are still infectious (4, 34). Similar studies of deletions of the shorter cytoplasmic domain of the Rous sarcoma virus envelope glycoprotein also demonstrated that this region was dispensable for virion incorporation, although in this case, infectivity was not affected (29). The studies reported herein, combined with systematic mutagenic analysis of the HIV-1 gp41 transmembrane region (11), essentially eliminate the possibility that interaction of specific gp41 sequences with gag proteins is a necessary step for incorporation of envelope glycoproteins into virions. The incorporation of the HIV-1 envelope glycoproteins into virions occurred more efficiently when levels of wild-type or mutant cell-associated envelope glycoproteins were low. This result suggests that a mechanism for concentrating the envelope glycoproteins into virions is involved in HIV-1 virion assembly. Whether the cytoplasmic domain is required for another step in virus assembly is not known, but the examination of HXB2(840-856) virions by electron microscopy and determination of the rate of gag precursor processing (Fig. 2D and E) did not reveal any gross defects in virus assembly or maturation (11a).

An HIV-1 provirus containing a 4-amino-acid deletion and an additional 15 random residues in the gp41 cytoplasmic domain, X10-1, was reported to result in a replication-competent HIV-1 exhibiting the ability to form syncytia but no ability to efficiently lyse single cells (9). A subsequent study also proposed a role for the cytoplasmic domain of gp41 in viral cytopathic effects (26). Contrary to these earlier studies, Kowalski et al. demonstrated that the replication and cytopathogenicity of the X10-1 mutant in Jurkat T lymphocytes were indistinguishable from those of the wild-type virus (23). The results of the present study provide further evidence that the cytoplasmic domain of gp41 does not play a major role in viral cytopathic effect independent of the effects of changes in this region on viral replication rate.

#### ACKNOWLEDGMENTS

We thank Kitty Lawrence for expert technical assistance, the AIDS Research and Reference Reagent Program and Bruce Walker for reagents, and William Haseltine for helpful discussions.

This work was supported by grants from the Leukemia Society of America (to J.S.) and the National Institutes of Health (AI01017, AI24755, AI28193, and AI07386). We also acknowledge the Center for AIDS Research grants (AI017386 and AI28691) for supporting necessary core facilities and for providing salary support for D.G.

#### REFERENCES

1. Andreassen, A., H. Bohr, J. Bohr, S. Brunak, T. Bugge, R. M. J. Cotterill, C. Jacobsen, P. Kusik, B. Lautrup, S. B. Petersen, T. Saermark, and K. Ulrich. 1990. Analysis of the secondary structure of the human immunodeficiency virus (HIV) proteins p17, gp120, and gp41 by computer modeling based on neural network methods. *J. Acquired Immune Defic. Syndr.* 3:615-622.
2. Berman, P. W., W. M. Nunes, and O. K. Haffar. 1988. Expression of membrane-associated and secreted variants of gp160 of human immunodeficiency virus type 1 in vitro and in continuous cell lines. *J. Virol.* 62:3135-3142.
3. Bordier, C. 1981. Phase separation of integral membrane proteins in Triton X-114 solution. *J. Biol. Chem.* 256:1604-1607.
4. Chakrabarti, L., M. Emerman, P. Tiollais, and P. Sonigo. 1989. The cytoplasmic domain of simian immunodeficiency virus transmembrane protein modulates infectivity. *J. Virol.* 63:4395-4403.
5. Cullen, B. R. 1987. Use of eukaryotic expression technology in the functional analysis of cloned genes. *Methods Enzymol.* 152:684-703.
6. Dalgleish, A. G., P. C. L. Beverly, P. R. Clapham, D. H. Crawford, M. F. Greaves, and R. A. Weiss. 1984. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature (London)* 312:763-766.
7. Earl, P. L., S. Koenig, and B. Moss. 1991. Biological and immunological properties of human immunodeficiency virus type 1 envelope glycoprotein: analysis of proteins with truncations and deletions expressed by recombinant vaccinia viruses. *J. Virol.* 65:31-41.
8. Earl, P. L., B. Moss, and R. W. Doms. 1991. Folding, interaction with GRP78-BiP, assembly, and transport of the human immunodeficiency virus type 1 envelope protein. *J. Virol.* 65:2047-2055.
9. Fisher, A. G., L. Ratner, H. Mitsuya, L. Marselle, M. E. Harper, S. Broder, R. C. Gallo, and F. Wong-Staal. 1986. Infectious mutants of HTLV-III with changes in the 3' region and markedly reduced cytopathic effects. *Science* 233:655-659.
10. Franchini, G., C. Gurgio, H. G. Guo, R. C. Gallo, E. Collati, K. A. Gargnoli, L. F. Hall, F. Wong-Staal, and M. S. Reitz. 1987. Sequence of simian immunodeficiency virus and its relationship to the human immunodeficiency viruses. *Nature (London)* 328:539-543.
11. Gabuzda, D., U. Olshevsky, P. Bertani, W. A. Haseltine, and J. Sodroski. 1991. Identification of membrane anchorage domains of the HIV-1 gp160 envelope glycoprotein precursor. *J. Acquired Immune Defic. Syndr.* 4:34-40.
- 11a. Gabuzda, D., and J. Sodroski. Unpublished data.
12. Gallaher, W. R., J. M. Ball, R. F. Garry, M. C. Griffin, and R. C. Montelaro. 1989. A general model for the transmembrane proteins of HIV and other retroviruses. *AIDS Res. Hum. Retroviruses* 5:431-440.
13. Grewe, C., A. Beck, and H. R. Gelderblom. 1990. HIV: early virus-cell interactions. *J. Acquired Immune Defic. Syndr.* 3:965-974.
14. Haffar, O. K., D. J. Dowbenko, and P. W. Berman. 1988. Topogenic analysis of the human immunodeficiency virus type 1 envelope glycoprotein, gp160, in microsomal membranes. *J. Cell Biol.* 107:1677-1687.
15. Haffar, O. K., D. J. Dowbenko, and P. W. Berman. 1991. The cytoplasmic tail of HIV-1 gp160 contains regions that associate with cellular membranes. *Virology* 180:439-441.
16. Haffar, O. K., G. R. Nakamura, and P. W. Berman. 1990. The carboxy terminus of human immunodeficiency virus type 1 gp160 limits its proteolytic processing and transport in transfected cell lines. *J. Virol.* 64:3100-3103.
17. Helseth, E., M. Kowalski, D. Gabuzda, U. Olshevsky, W. Haseltine, and J. Sodroski. 1990. Rapid complementation assays measuring replicative potential of human immunodeficiency virus type 1 envelope glycoprotein mutants. *J. Virol.* 64:2416-2420.
18. Helseth, E., U. Olshevsky, D. Gabuzda, B. Ardman, W. Haseltine, and J. Sodroski. 1990. Changes in the transmembrane region of the human immunodeficiency virus type 1 gp41 enve-



- lope glycoprotein affect membrane fusion. *J. Virol.* 64:6314-6318.
19. Hirsch, V., N. Riedel, and J. L. Mullins. 1987. Genome organization of STLV-3 is similar to that of the AIDS virus except for a truncated transmembrane protein. *Cell* 49:307-319.
  20. Hirsch, V. A., P. Edmondson, J. Murphey-Corb, B. Arbeille, P. R. Johnson, and J. L. Mullins. 1989. SIV adaptation to human cells. *Nature (London)* 341:573-574.
  21. Klatzmann, D., E. Champagne, S. Chamaret, J. Gruest, D. Guetard, T. Hercent, J. C. Gluckman, and L. Montagnier. 1984. T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature (London)* 312:767-768.
  22. Kodama, T., D. Wooley, Y. Naidu, H. W. Kestler, M. D. Daniel, Y. Li, and R. Desrosiers. 1989. Significance of premature stop codons in *env* of simian immunodeficiency virus. *J. Virol.* 63:4709-4714.
  23. Kowalski, M., L. Bergeron, T. Dorfman, W. Haseltine, and J. Sodroski. 1991. Attenuation of human immunodeficiency virus type 1 cytopathic effect by a mutation affecting the transmembrane envelope glycoprotein. *J. Virol.* 65:281-291.
  24. Kowalski, M., J. Potz, L. Basiripour, T. Dorfman, W. C. Goh, E. Terwilliger, A. Dayton, C. Rosen, W. Haseltine, and J. Sodroski. 1987. Functional regions of the envelope glycoprotein of human immunodeficiency virus type 1. *Science* 237:1351-1355.
  25. Kunkel, T. A., J. D. Roberts, and R. A. Zakour. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* 154:367-382.
  26. Lee, S. J., W. Hu, A. Fisher, D. Looney, V. Kan, H. Mitsuya, L. Ratner, and F. Wong-Staal. 1989. Role of the carboxyl-terminal portion of the HIV-1 transmembrane protein in viral transmission and cytopathogenicity. *AIDS Res. Hum. Retroviruses* 5:441-449.
  27. McDougal, J. S., M. Kennedy, J. Sligh, S. Cort, A. Mowle, and J. Nicholson. 1986. Binding of the HTLV-III/LAV to T4+ T cells by a complex of the 110 K viral protein and the T4 molecule. *Science* 231:382-385.
  28. Myers, G., S. F. Josephs, A. B. Rabson, T. I. Smith, and F. Wong-Staal (ed.). 1988. Human retroviruses and AIDS. Los Alamos National Laboratory, Los Alamos, N.Mex.
  29. Perez, L. G., G. L. Davis, and E. Hunter. 1987. Mutants of the Rous sarcoma virus envelope glycoprotein that lack the transmembrane anchor and cytoplasmic domains: analysis of intracellular transport and assembly into virions. *J. Virol.* 61:2981-2988.
  30. Queen, C., and D. Baltimore. 1983. Immunoglobulin gene transcription is activated by downstream sequence elements. *Cell* 33:741-748.
  31. Ratner, L., A. Fisher, L. L. Jagodzinski, H. Mitsuya, R. S. Lion, R. C. Gallo, and F. Wong-Staal. 1987. Complete nucleotide sequence of functional clones of the AIDS virus. *AIDS Res. Hum. Retroviruses* 3:57-69.
  32. Ratner, L., W. Haseltine, R. Patarca, K. J. Livak, B. Starcich, S. F. Josephs, E. R. Doran, J. A. Rafalski, E. A. Whitehorn, K. Baumeister, L. Ivanoff, S. R. Petteway, M. L. Pearson, J. A. Lautenberger, T. S. Papas, J. Ghayeh, N. T. Chang, R. C. Gallo, and F. Wong-Staal. 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature (London)* 313:277-284.
  33. Rho, H. M., B. Poiesz, W. Ruscetti, and R. C. Gallo. 1981. Characterization of the reverse transcriptase from a new retrovirus (HTLV) produced by a human cutaneous T-cell lymphoma cell line. *Virology* 112:355-360.
  34. Rice, N. R., L. E. Henderson, R. C. Sowder, T. D. Copeland, S. Oroszlan, and J. F. Edwards. 1990. Synthesis and processing of the transmembrane envelope protein of equine infectious anemia virus. *J. Virol.* 64:3770-3778.
  35. Stein, B. S., S. D. Gouda, J. D. Lifson, R. C. Penhallow, K. G. Bensch, and E. G. Engleman. 1987. pH-independent HIV entry into CD4-positive T cells via virus envelope fusion to the plasma membrane. *Cell* 49:659-668.
  36. Terwilliger, E., B. Godin, J. Sodroski, and W. Haseltine. 1989. Construction and use of replication-competent human immunodeficiency virus that expresses the CAT enzyme. *Proc. Natl. Acad. Sci. USA* 86:3857-3861.
  37. Terwilliger, E., J. Sodroski, C. Rosen, and W. Haseltine. 1986. Effects of mutations within the 3' *orf* open reading frame region of human T-cell lymphotropic virus type III (HTLV-III/LAV) on replication and cytopathogenicity. *J. Virol.* 60:754-760.
  38. Thali, M., U. Olshevsky, C. Furman, D. Gabuzda, J. Li, and J. Sodroski. 1991. Effects of changes in gp120-CD4 binding affinity on human immunodeficiency virus type 1 envelope glycoprotein function and soluble CD4 sensitivity. *J. Virol.* 65:5007-5012.
  39. Venable, R. M., R. W. Pastor, B. R. Brooks, and F. W. Carson. 1989. Theoretically determined three-dimensional structures for amphipathic segments of the HIV-1 gp41 envelope protein. *AIDS Res. Hum. Retroviruses* 5:7-22.
  40. Willey, R. L., J. S. Bonafacino, B. J. Potts, M. A. Martin, and R. D. Klausner. 1988. Biosynthesis, cleavage, and degradation of the human immunodeficiency virus 1 envelope glycoprotein gp160. *Proc. Natl. Acad. Sci. USA* 85:9580-9584.
  41. Zack, J. A., S. J. Arrigo, S. R. Weitsman, A. S. Go, A. Haislip, and I. S. Y. Chen. 1990. HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure. *Cell* 61:213-222.

## Defects in functional expression of an influenza virus hemagglutinin lacking the signal peptide sequences

(vectorial discharge/polypeptide transport/glycosylation/hemagglutination)

KENJI SEKIKAWA AND CHING-JUH LAI

Molecular Viral Biology Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20205

Communicated by Robert M. Chanock, March 7, 1983

**ABSTRACT** We have investigated the requirement of the signal sequence for expression of influenza virus hemagglutinin (HA). For this purpose we used a recombinant prepared from a late-region deletion mutant of simian virus 40 (SV40) and cloned influenza HA DNA; the influenza DNA was inserted into the late region of SV40 previously occupied by the deleted sequences coding for SV40 capsid proteins. A simple in-phase deletion was made in the HA DNA, resulting in loss of 11 internal amino acids from the 16 amino acid signal peptide. This deletion HA recombinant was then used to infect African green monkey kidney cells. Mutant HA was not detected on the cell surface but stably accumulated in the cytoplasm at a level similar to that of wild-type HA. NaDodSO<sub>4</sub>/polyacrylamide gel analysis of lysates from infected cells showed that mutant HA was not glycosylated. Significantly, the amount of mutant HA synthesized was not affected by tunicamycin. In contrast, wild-type HA was decreased more than 90% by tunicamycin. These findings suggest that mutant polypeptide is synthesized on free polyribosomes rather than on membrane-bound polyribosomes. The mutant HA failed to agglutinate erythrocytes, probably due to a defect directly or indirectly associated with the lack of carbohydrate side chains.

The hemagglutinin (HA) of influenza virus constitutes the major viral envelope glycoprotein and is responsible for binding of virus to sialic acid receptors on the cell surface (1). HA also mediates membrane fusion that initiates uncoating of the virion inside the infected cell (2-5). During infection HA is synthesized in the cytoplasm, glycosylated, and transferred to the outer cell membrane where the final stage of virus assembly takes place. Functional HA contains three HA polypeptide subunits and a specific protease cleavage of each subunit is required for viral infectivity (6, 7).

Sequence analysis of HA from many influenza virus strains shows the presence of a hydrophobic sequence at the amino terminus and another hydrophobic sequence near the carboxyl terminus (8-11). The hydrophobic carboxyl terminus is responsible for anchoring the polypeptide on the outer membrane because HA lacking these sequences is secreted extracellularly (12). On the other hand, the hydrophobic amino terminus is cleaved during maturation and is therefore absent in the membrane-integrated HA (13). These transient hydrophobic sequences, commonly referred to as the signal peptide, exist in all known eukaryotic transmembrane and secretory proteins with several notable exceptions, such as ovalbumin and the influenza virus neuraminidase (14-16). Similar transient signal peptides also exist in the periplasmic proteins of prokaryotes. It has been postulated that the signal peptide segments play a crucial role in mediating the transfer of polypeptide across the membrane by a specific mechanism known as

vectorial discharge of nascent polypeptide (17, 18). Synthesis and transfer of proteins utilize a specialized machinery of membrane-bound polyribosomes and the Golgi apparatus in eukaryotes. Mutational analysis of the signal sequences to further elucidate the transfer process has not been reported in eukaryotic systems.

Previously we constructed a recombinant consisting of a late-region deletion mutant of simian virus 40 (SV40) and cloned influenza DNA representing the complete sequences of the gene coding for the HA surface glycoprotein. The cloned influenza DNA was inserted into the late region of SV40 previously occupied by the deleted sequences coding for SV40 capsid proteins. This HA-SV40 recombinant expressed fully functional HA in African green monkey kidney (AGMK) cells. The HA was glycosylated and inserted into the outer membrane of the infected cell (19). This system was used in the present study to examine the role of the signal peptide in HA maturation and function. In this study, we describe the construction of a deletion recombinant that produced HA lacking the signal peptide sequences. Analyses of functional defects in the mutant HA that affect glycosylation, maturation at the cell surface, and hemagglutination activity are presented.

### MATERIALS AND METHODS

**Preparation of DNA and Construction of HA-SV40 Deletion Mutants.** Previously we cloned a full-length DNA copy coding for the HA of influenza virus strain A/Udm/72 (H3N2). To achieve functional expression the cloned HA DNA was inserted into the late region of a SV40 vector and a recombinant of HA-SV40 that produced a functionally active HA was isolated. Subsequently the HA-SV40 DNA recombinant was cloned in pBR322 by using the unique *Bam*HI cleavage site (19).

The recombinant pHA-SV40 DNA was prepared from a cloned *Escherichia coli* transformant as described (19). DNA fragments prepared by restriction enzyme digestion were separated on 3.5% polyacrylamide gels and eluted electrophoretically. The *Mbo* II DNA fragment downstream of the cleavage site was digested with  $\lambda$  exonuclease (28 units/ml) in a buffer containing 67 mM glycine (pH 9.4), 3 mM MgCl<sub>2</sub>, 3 mM 2-mercaptoethanol at 0°C for 30 min (20). Other enzyme digestions were carried out according to the prescribed conditions by their suppliers. Constructed plasmid recombinants carrying a signal peptide deletion were subjected to sequence analysis by the method of Maxam and Gilbert (21). Mutant HA-SV40 DNA was isolated after *Bam*HI digestion, circularized, and used for coinfection of AGMK cells with a tsA28 SV40 early-region mutant helper virus.

**Radiolabeling and Analysis of HA Polypeptides.** AGMK cells infected with the mutant HA-SV40 and incubated at 40°C for

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: HA, hemagglutinin; SV40, simian virus 40; AGMK, African green monkey kidney.

72 hr were labeled with [ $^{35}$ S]methionine (100  $\mu$ Ci/ml; 1 Ci =  $3.7 \times 10^{10}$  Bq) in a methionine-free medium for 5 hr. Labeling of cells during infection with the wild-type HA-SV40 was carried out similarly. Cells were directly lysed in radioimmuno-precipitation assay buffer and the supernatant fraction was immunoprecipitated for analysis on NaDodSO<sub>4</sub>/polyacrylamide gels (19).

**Indirect Immunofluorescence Assay.** Three days after infection with either wild-type or mutant HA-SV40, unfixed cell monolayers were stained directly in a live cell fluorescence assay with sheep antiserum against HA and fluorescein-conjugated rabbit anti-sheep IgG. Fixed cell fluorescence assay was performed on cells treated with 3.7% formaldehyde for 10 min and 0.1% Triton X-100 for 5 min as described by Wehland *et al.* (22).

**Hemagglutination Test.** Lysates of recombinant virus-infected cells were prepared in phosphate-buffered saline and serially diluted in a microtiter plate. Guinea pig erythrocytes (0.1% suspension in phosphate-buffered saline) were added to each dilution and hemagglutination activity was examined after 30 min at room temperature.

## RESULTS

**Construction of Mutant HA Encoding Deleted Signal Sequences.** The wild-type HA-SV40 recombinant cloned previously at the unique *Bam*HI site of pBR322 was used for the derivation of HA deletion mutants. Fig. 1 illustrates the strategy used to construct HA mutants lacking the signal peptide sequences. *Endo*R *Mbo* II cleaves the HA-SV40 plasmid once at a site 12 base pairs downstream of the initiation codon of the HA and several other sites outside the HA gene. Digestion of pHA-SV40 with *Mbo* II yielded two DNA fragments containing HA-specific sequences. The DNA fragment downstream of the cleavage site was briefly treated with  $\lambda$  exonuclease to shorten both DNA strands progressively, whereas the other *Mbo* II

fragment was not treated. After nuclease S1 digestion, both HA DNA fragments were joined. In this manner, a series of deletion mutations were introduced at the *Mbo* II cleavage site while retaining the upstream *Mbo* II enzyme recognition site. The joined HA DNA from the ligase reaction mixture was cleaved with *Kpn* I and *Xba* I to generate cohesive termini. This specific DNA fragment was isolated and rebuilt into the pHA-SV40 plasmid by joining it to the other fragment from the *Kpn* I/*Xba* I digest of pHA-SV40. In this manner, the signal deletion was introduced into the wild-type HA DNA molecule and represented the only mutation in the HA-SV40 recombinant. The constructed DNA was used directly for cloning by transformation of *E. coli*.

**Sequence Analysis of Mutant HA Lacking the Signal Peptide.** The constructed deletion mutant DNAs were isolated from *E. coli* transformants and characterized by digestion with *Endo*R *Ava* II and polyacrylamide gel electrophoresis. For comparison, similar *Ava* II digestion of the wild-type pHA-SV40 yielded a 338-nucleotide fragment containing the signal sequences and this type of analysis allowed us to estimate the size of the corresponding fragments of mutant DNA. These estimates allowed us to determine the extent of deletion and served to guide us in choosing mutant recombinants for sequence analysis. Among six mutant isolates shortened by 30–100 nucleotides, we found one mutant that contained an in-phase deletion, whereas the other five showed either a +1 or a –1 frameshift deletion. The latter were not analyzed further. The wild-type nucleotide sequence and the encoded signal peptide sequence at the amino terminus of HA were compared with the sequences of the HA-SV40 mutant (Fig. 2). The HA signal peptide consists of 16 amino acids—1 charged amino acid (Lys), 11 hydrophobic amino acids (Met, Ile, Ala, Leu, Phe, Val, and Gly), and 4 uncharged polar amino acids (Thr, Ser, Tyr, and Cys). The proteolytic cleavage site is located at the Gly-Gln juncture because Gln is the amino terminus of the membrane-incorporated HA molecules. The mutant HA sustained a deletion of 33 base pairs or 11 amino acids, all within the signal peptide. As a result, the remaining signal region contains three hydrophobic amino acids (Met, Ile, and Gly), a polar uncharged amino acid (Thr), and a charged amino acid (Lys). The signal peptide cleavage site of Gly-Gln and the subsequent sequences remained unchanged in the mutant.

**Analysis of HA by Indirect Immunofluorescence Assay.** An indirect immunofluorescence assay was employed to identify the intracellular site of accumulation of the mutant HA in the infected cell. Initially, the deletion mutant genome of HA-SV40 was packaged within SV40 virions and propagated together with the SV40 *ts* helper virus at the restrictive temperature (40°C). AGMK cells were infected with the SV40 *ts* helper and either the mutant or the wild-type HA-SV40 recombinant for 3 days prior to assay for HA by live cell fluorescence staining with HA antibody and an appropriate fluorescein conjugate. Fluorescence was not detected on mutant infected cells, whereas cells infected with wild-type HA-SV40 showed positive surface staining (data not shown). When infected cells were fixed and stained for HA fluorescence, wild-type HA was detected in the cytoplasm and on the membrane. In contrast, the mutant HA was found only in the cytoplasm. The mutant HA was stained as intensively as the wild-type HA (Fig. 3). These results demonstrate that mutant HA lacking the signal sequences accumulates in the cytoplasm but is not incorporated into the outer membrane.

**Analysis of HA by NaDodSO<sub>4</sub>/Polyacrylamide Gel Electrophoresis.** We estimated the molecular size of the mutant HA product to determine whether the mutant polypeptide is post-translationally modified by glycosylation in the same manner as

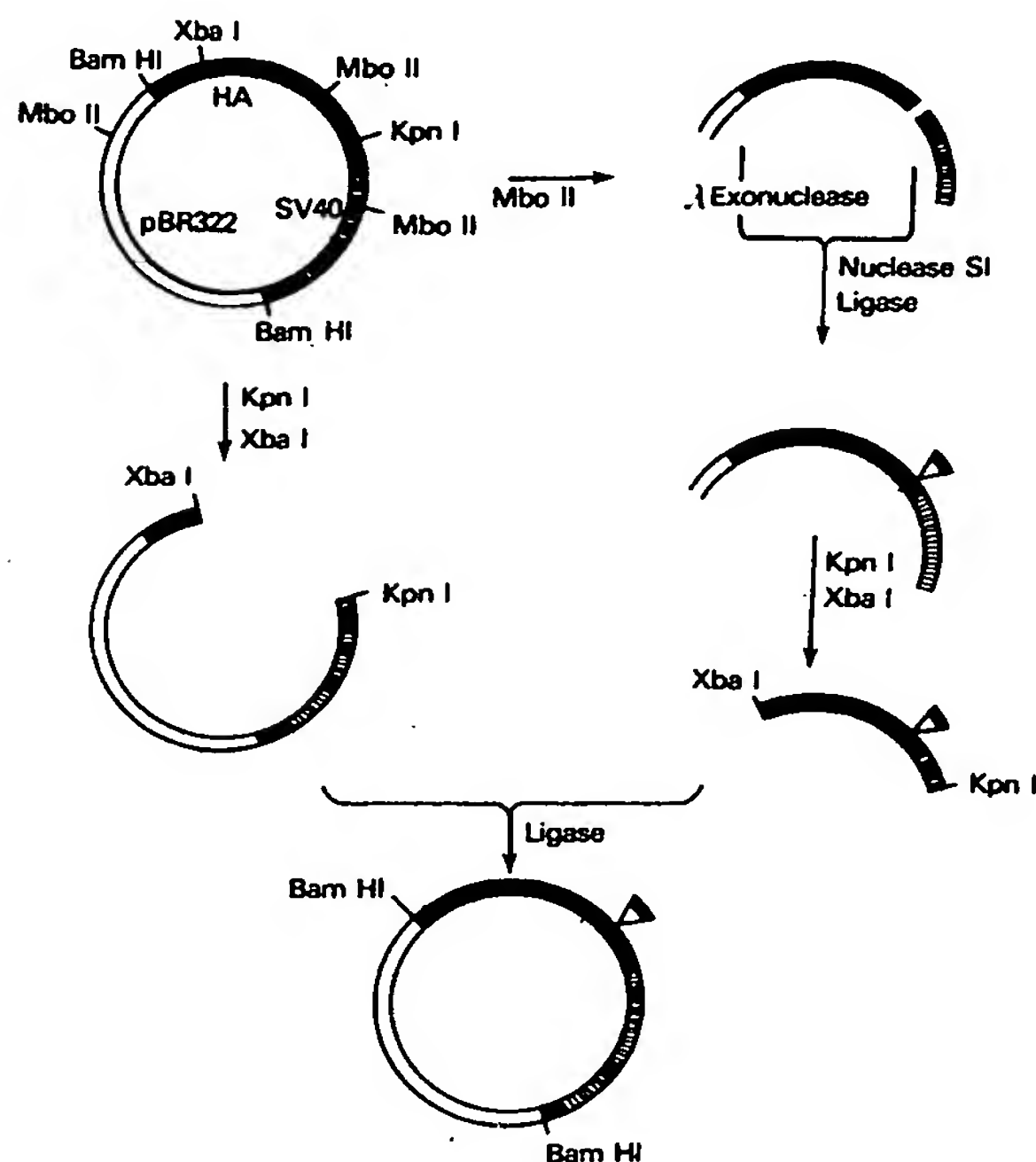
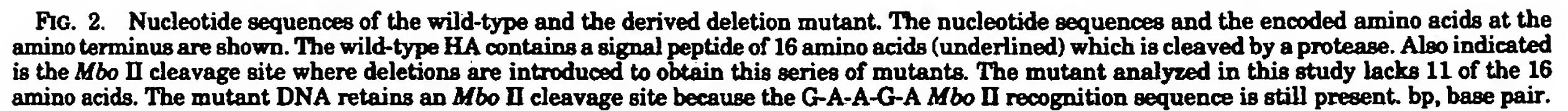


FIG. 1. Construction of signal peptide deletion mutants of HA-SV40.





tons, similar in size to the glycosylated HA produced during influenza virus infection (Fig. 4). In the presence of tunicamycin, an inhibitor that prevents polypeptide glycosylation (23), this value was decreased to 56,000 daltons, the predicted size for unglycosylated HA. It should be noted that addition of tunicamycin also decreased the accumulation of unglycosylated wild-type HA by  $\approx 90\%$ . Similar analysis showed that the mutant recombinant produced a HA polypeptide of 56,000 daltons equivalent to the predicted molecular size for unglycosylated HA. When tunicamycin was added during radiolabeling, the mutant HA product that corresponded to the most prominent band was not decreased in size and remained at 56,000 daltons. Other minor bands from the deletion mutant were probably derived from rearranged HA sequences that arose during passages in AGMK cells. These results provide evidence that the mutant polypeptide synthesized during infection was not modified by glycosylation posttranslationally. In contrast to the wild-type HA tunicamycin did not show an inhibitory effect on the synthesis of mutant HA.

Western blot analysis showing the effect of tunicamycin on the expression of HA-tagged proteins. The blot is divided into three main sections: **M** (Molecular Weight Marker), **wt-HA** (wild-type HA-tagged protein), and **dl-HA** (deleted HA-tagged protein). Each section has two lanes: **-** (without tunicamycin) and **+** (with tunicamycin). The **Mock** section shows a control without any HA-tagged protein. Molecular weight markers are indicated on the left: 70-75 (HAo), 56, 28, and 26 kDa. The **wt-HA** section shows a prominent band at 70-75 kDa in the **-** lane, which is significantly reduced in the **+** lane. The **dl-HA** section shows a prominent band at 56 kDa in both the **-** and **+** lanes. The **Mock** section shows no bands. The **M** section shows bands at 28 and 26 kDa.

FIG. 4. NaDodSO<sub>4</sub>/polyacrylamide gel analysis of the wild-type and mutant HAs. Monolayers of AGMK cells were infected with the wild-type HA-SV40 (wt-HA) and mutant HA-SV40 (dl-HA) for 72 hr. Infected cells were pretreated with tunicamycin (1 μg/ml) for 60 min prior to labeling with [<sup>35</sup>S]methionine (100 μCi/ml) in methionine-free medium in the presence of 1 μg of tunicamycin per ml. Cell lysates were immunoprecipitated with HA antiserum and analyzed on 15% NaDodSO<sub>4</sub>/polyacrylamide gels. M indicates the polypeptide markers (shown in daltons × 10<sup>-3</sup>) prepared from [<sup>35</sup>S]methionine-labeled influenza virions.

either the wild-type HA-SV40 or the mutant HA-SV40 and were examined for ability to agglutinate guinea pig erythrocytes. In a control experiment, a hemagglutination titer of 1:8 was observed for the wild-type HA. The lysate containing the unglycosylated mutant HA at an antigen concentration comparable to the wild-type HA did not show detectable hemagglutination activity (data not shown).

## DISCUSSION

This study concerns the functional role of HA signal sequences in the processing and maturation of the H3 subtype HA at the cell surface. We constructed deletion mutants containing a deletion of the signal peptide at the *Mbo* II cleavage site. *Mbo* II cleaves at a site eight nucleotides downstream from its recognition sequence 5' G-A-A-G-A 3' such that a one-nucleotide 3' protrusion is generated. The current construction scheme allowed the retention of sequences including the *Mbo* II recognition site and those encoding the first three amino acids. After construction the deletion juncture should become a new *Mbo* II cleavage site; however, in our mutant, two additional bases were deleted from the upstream fragment for reasons not clear to us. This created a new *Mbo* II cleavage site three bases after the juncture. Nonetheless, it should be possible to apply the scheme used in our study for the construction of deletions of any length and thereby obtain a series of progressively extended HA signal peptide mutants. Analysis of these HA mutants should yield information concerning the signal sequences required for functional activity. The mutant constructed in this study, with a deletion of 11 of the 16 amino acids in the signal sequence, was suitable for characterization of defective functions. This mutant retained the presumptive amino-terminal signal-peptide protease cleavage site and therefore specific cleavage at this site did not preclude the observed abnormalities in HA function.

Our results demonstrate that the deleted HA polypeptide from the constructed mutant is not posttranslationally modified by glycosylation despite the presence of all normal Asn-X-Thr and Asn-X-Ser sequences. The signal hypothesis predicts that wild-type HA mRNA is translated on membrane-bound polyribosomes of the endoplasmic reticulum and the nascent HA polypeptides are subsequently transferred to the lumen by a mechanism of translocation and vectorial discharge. There is considerable evidence suggesting that glycosylation takes place initially on nascent polypeptides on endoplasmic reticulum membranes and subsequently the carbohydrate components are modified in the Golgi apparatus (24, 25). It is interesting to note that production of wild-type, fully glycosylated HA (in the absence of tunicamycin) was 10-fold higher than that of unglycosylated HA (in the presence of tunicamycin). One possible explanation is that tunicamycin inhibition freezes these glycosylation steps and hence translation of HA mRNA is slowed down, although a low level synthesis of unglycosylated wild-type HA is eventually completed. Alternatively, tunicamycin increases the breakdown of unglycosylated wild-type HA because glycosylation confers intracellular stability and resistance to proteolytic digestion (26). On the other hand, translation of mutant HA mRNA probably proceeds on free polyribosomes and therefore the nascent polypeptide lacking the signal sequences fails to translocate. The nascent as well as the completed mutant polypeptides do not encounter glycosyl transferases, which are membrane-associated, and the unglycosylated HA molecules remain in the cytoplasm. The finding that tunicamycin exerts no detectable inhibition on the accumulation of mutant HA further supports the model of synthesis on free polyribosomes.

Our results show that the HA polypeptide lacking the tran-

sient signal sequences is not detectable at the cell surface. The transport of polypeptides from the cytoplasm to the outer membranes or for extracellular secretion is a complex process, presumably initiated by the signal sequences. Recent evidence suggests that the signal sequences at the amino terminus of polypeptides specify the initial recognition by a multiple-component structure known as signal peptide recognition particles (27, 28). This recognition process did not occur with the mutant HA polypeptide in simian cells. In *E. coli*, similar hydrophobic sequences that provide a transient signal for polypeptide transport also exist in several periplasmic proteins. Deletion or alteration of these sequences results in accumulation of these proteins in the cytoplasm (29, 30).

The HA-SV40 mutant produces unglycosylated HA that is functionally inactive, as assayed by the hemagglutination test. This failure to agglutinate erythrocytes suggests at least two possibilities. First, the mutant HA may be monomeric and thus fail to assemble into a trimeric structure, which is required for functional activity. Second, the defect of mutant HA may be due to the lack of carbohydrate components and as a result the mutant HA subunit fails to assemble properly. Because 50% of the carbohydrate side chains of HA can be removed without affecting hemagglutination of infectivity (31), it would seem that at least some carbohydrate components are not crucial. Therefore, it is most likely that the assembly process to form mature active surface HA may require a defined configuration of the polypeptide subunit that is programmed by posttranslational modifications such as glycosylation.

Other observations support the view that deletion within the signal peptide of the HA polypeptide is most likely responsible for the functional defects described in this paper. Recently we have found that several HAs containing substitutions of amino acids in the signal peptide produced by point-mutation in HA-SV40 also showed similar defects of glycosylation and cell surface expression (unpublished data). This finding effectively rules out the possibility that a mutation other than the signal sequence deletion in our mutant was responsible for the observed functional abnormalities.

There are a number of differences evident when our results are compared with those of a similar recently published study (32). First, we constructed mutants that contained a simple internal deletion in the HA signal sequence. The deletion HA that was studied extensively preserved the first four amino-terminal amino acids and the remaining sequences. Therefore, characterization of this mutant allowed us to correlate specific functional defects with this deletion. In contrast, using a different construction procedure, other investigators derived a mutant that sustained a deletion in the 5'-noncoding region of the HA recombinant. In addition, three different amino-terminal amino acids were added to replace the entire signal peptide (except the carboxyl-terminal Gly). Presumably as a consequence of either or both of these alterations, the mutant HA was synthesized at a level 0.16% that of wild-type HA. Hence, it is difficult to interpret the lack of glycosylation of the small number of HA molecules that were produced. In contrast, mutant HA polypeptides directed by our mutant were made in the amount similar to that of wild-type HA molecules. Consequently, the absence of glycosylation of our mutant HA and its failure to migrate to the cell membrane can be convincingly attributed to deletion within the signal peptide.

The authors thank Dr. Robert M. Chanock for helpful discussions and useful suggestions about the manuscript. We also thank Ms. Jo Ann Berndt and Ms. Salome Kruger for their excellent technical assistance and Ms. Linda Jordan for typing the manuscript.

1. Hirst, G. K. (1942) *J. Exp. Med.* 75, 49-64.
2. Laver, W. G. & Valentine, R. C. (1969) *Virology* 38, 105-119.
3. Wiley, D. C., Skehel, J. J. & Waterfield, M. (1977) *Virology* 79, 446-448.
4. Matlin, K., Reggio, H., Helenius, A. & Simons, K. (1981) *J. Cell Biol.* 91, 601-613.
5. Huang, R. T. C., Rott, R., Wahn, K., Klenk, H.-D. & Kahama, T. (1980) *Virology* 107, 313-319.
6. Lazarowitz, S. G. & Choppin, P. W. (1975) *Virology* 68, 440-454.
7. Klenk, H. D., Rott, R., Orlich, M. & Blodorn, J. (1975) *Virology* 68, 426-439.
8. Porter, A. G., Barber, C., Carey, N. H., Hallewell, R. A., Threlfall, G. & Emtage, J. S. (1979) *Nature (London)* 282, 471-477.
9. Gething, M.-J., Bye, J., Skehel, J. & Waterfield, M. (1980) *Nature (London)* 287, 301-306.
10. Min Jou, W., Verhoeyen, M., Devos, R., Saman, E., Fang, R., Huylebroeck, D., Fiers, W., Threlfall, G., Barber, C., Carey, N. & Emtage, S. (1980) *Cell* 19, 683-696.
11. Air, G. M. (1981) *Proc. Natl. Acad. Sci. USA* 78, 7639-7643.
12. Sveda, M. M., Markoff, L. J. & Lai, C.-J. (1982) *Cell* 30, 649-659.
13. McCauley, J., Skehel, J., Elder, K., Gething, M.-J., Smith, A. & Waterfield, M. (1980) in *Structure and Variation in Influenza Virus*, eds. Laver, G. & Air, G. (Elsevier/North-Holland, New York), pp. 97-105.
14. Sabatini, D. D., Kreibich, G., Morimoto, T. & Adesnik, M. (1982) *J. Cell Biol.* 92, 1-22.
15. Fields, S., Winter, G. & Brownlee, G. G. (1981) *Nature (London)* 290, 213-217.
16. Davis, B. D. & Tai, P.-C. (1980) *Nature (London)* 283, 433-438.
17. Blobel, G. & Dobberstein, B. (1975) *J. Cell Biol.* 67, 835-851.
18. Milstein, C., Brownlee, G. G., Harrison, T. M. & Mathews, M. B. (1972) *Nature (London) New Biol.* 239, 117-120.
19. Sveda, M. M. & Lai, C.-J. (1981) *Proc. Natl. Acad. Sci. USA* 78, 5488-5492.
20. Carbon, J., Shenk, T. E. & Berg, P. (1975) *Proc. Natl. Acad. Sci. USA* 72, 1392-1396.
21. Maxam, A. H. & Gilbert, W. (1980) *Methods Enzymol.* 65, 499-560.
22. Wehland, J., Willingham, M. C., Gall, M. G. & Pastan, I. (1982) *Cell* 28, 831-841.
23. Takatsuki, A., Kohno, K. & Tamura, A. (1975) *Agric. Biol. Chem.* 39, 2089-2091.
24. Rothman, J. E., Katz, F. N. & Lodish, H. F. (1978) *Cell* 15, 1447-1454.
25. Rothman, J. E. & Fine, R. E. (1980) *Proc. Natl. Acad. Sci. USA* 77, 780-784.
26. Sibley, C. H. & Wagner, R. A. (1981) *J. Immunol.* 126, 1868-1873.
27. Meyer, D. I., Krause, E. & Dobberstein, B. (1982) *Nature (London)* 297, 647-650.
28. Walter, P. & Blobel, G. (1982) *Nature (London)* 299, 691-698.
29. Bassford, P. J. & Beckwith, J. (1979) *Nature (London)* 277, 538-541.
30. Inouye, S., Soberon, X., Franceschini, T., Nakamura, K., Itakura, K. & Inouye, M. (1982) *Proc. Natl. Acad. Sci. USA* 79, 3438-3441.
31. Collins, J. K. & Knight, C. A. (1978) *J. Virol.* 27, 164-171.
32. Gething, M.-J. & Sambrook, J. (1982) *Nature (London)* 300, 598-603.



## CD28 Interaction with B7 Costimulates Primary Allogeneic Proliferative Responses and Cytotoxicity Mediated by Small, Resting T Lymphocytes

By Miyuki Azuma,\* Mark Cayabyab,\* David Buck,† Joseph H. Phillips,\* and Lewis L. Lanier\*

From the \*Department of Immunology, DNAX Research Institute for Molecular and Cellular Biology, Palo Alto, California 94304; and the †Research Department, Becton Dickinson Immunocytometry Systems, San Jose, California 95131

### Summary

Engagement of the CD3/T cell antigen receptor complex on small, resting T cells is insufficient to trigger cell-mediated cytotoxicity or to induce a proliferative response. In the present study, we have used genetic transfection to demonstrate that interaction of the B7-BB1 B cell activation antigen with the CD28 T cell differentiation antigen costimulates cell-mediated cytotoxicity and proliferation initiated by either anti-CD2 or anti-CD3 monoclonal antibody (mAb). Moreover, a B7-negative Burkitt's lymphoma cell line that fails to stimulate an allogeneic mixed lymphocyte response is rendered a potent stimulator after transfection with B7. The mixed leukocyte reaction proliferative response against the B7 transfectant is inhibited by either anti-CD28 or B7 mAb. We also demonstrate that freshly isolated small, resting human T cells can mediate anti-CD3 or anti-CD2 mAb-redredirected cytotoxicity against a murine Fc receptor-bearing mastocytoma transfected with human B7. These preexisting cytotoxic T lymphocytes in peripheral blood are present in both the CD4 and CD8 subsets, but are preferentially within the CD45RO<sup>+</sup> "memory" population. While small, resting T cells apparently require costimulation by CD28/B7 interactions, this requirement is lost after T cell activation. Anti-CD3 initiates a cytotoxic response mediated by in vitro cultured T cell clones in the absence of B7 ligand. The existence of functional cytolytic T cells in the small, resting T cell population may be advantageous in facilitating rapid responses to immune challenge.

**T** lymphocytes recognize antigen via the CD3/TCR complex. Binding of anti-CD3 or anti-TCR mAb to T cells results in a rapid increase in intracellular [Ca<sup>2+</sup>] and the generation of inositol triphosphate (IP<sub>3</sub>) (1-4). Although crosslinking of CD3/TCR alone is often sufficient to induce inositolphosphate pathway activation, other signals are apparently necessary to induce functions such as cytokine secretion and proliferation (4). Both soluble factors and interaction with cell surface receptors have been implicated as costimulators of CD3/TCR-mediated activation. For example, allo-antigen-specific helper T cells fail to produce cytokines in response to HLA-DR7 antigen expressed in murine L cells. However, cotransfection of HLA-DR7 and CD54 (intracellular adhesion molecule 1 [ICAM-1]), a cellular ligand for CD11/18 leukocyte function-associated molecule 1 [LFA-1]), restores T cell responsiveness (5). Thus, antigen-specific recognition alone may be insufficient to trigger effector functions since additional accessory molecules are required for an efficient response.

CD28 is a disulfide-linked homodimer that is expressed on the majority of human peripheral blood T cells (6, 7).

mAbs against CD28 in conjunction with phorbol ester induce T cell proliferation (7) and augment proliferation induced by anti-CD3 or anti-CD2 mAb (8-12). Anti-CD28 mAb-induced proliferation is IL-2 dependent (7, 8) and possibly results from stabilization of messenger RNA for IL-2 and other cytokines (13).

Recently, it has been demonstrated that a B cell activation antigen, B7 (14) or BB1 (15), is a natural ligand for CD28 and that this receptor/ligand interaction mediates heterotypic adhesion (16). Interaction of CD28 and B7 results in augmentation of T cell proliferation and cytokine production, and antibodies against CD28 or B7 inhibit alloantigen and mitogen-induced proliferative responses (16, 17). B7 is weakly expressed on resting B cells and monocytes, and is elevated after stimulation with pokeweed mitogen, anti-Ig, anti-HLA class II, and EBV (14, 15, 17). Therefore, it is likely that T cell activation via CD28 will depend on the capacity of the APC to upregulate expression of B7. The importance of the CD28/B7 interaction in polyclonal mitogenic responses prompted us to investigate whether the binding of B7 to CD28

costimulates the generation of a proliferative response against alloantigen and can augment the generation of T cell-mediated cytotoxicity in small, resting T lymphocytes.

## Materials and Methods

**Preparation of T Lymphocytes.** Human peripheral blood was obtained from the Stanford Blood Center (Palo Alto, CA). Mononuclear cells were isolated by density gradient centrifugation using Ficoll/Hypaque. Monocytes and B cells were depleted by plastic adherence and passage through nylon wool columns (18). Nonadherent lymphocytes were fractionated by centrifugation on discontinuous gradients consisting of 30% and 40% Percoll (Pharmacia Fine Chemicals, Piscataway, NJ) in PBS containing 10% FCS (18). High buoyant density lymphocytes enriched for small, resting T cells were isolated from the bottom of the Percoll gradients. Thymocytes from 1-yr-old children undergoing cardiac surgery were obtained from Stanford Medical Center (Palo Alto, CA). The biopsies were minced with scissors, the single cell suspension was passed through a nylon mesh, and mononuclear cells were isolated by Ficoll/Hypaque density centrifugation. Umbilical cord blood was collected from healthy full-term neonates immediately after vaginal delivery. High buoyant density cord blood CD3<sup>+</sup> T lymphocytes were obtained as for peripheral blood.

**T Cell Clones.** T cell clones were established from normal PBL or fetal liver, as described (19). The following clones were used: 1375b2.22 is a TCR- $\gamma/\delta$ <sup>+</sup> CTL clone established from fetal liver; 1360b.1 and 1320.3 are TCR- $\alpha/\beta$ <sup>+</sup> CTL clones established from fetal liver. Jp28a.6 is a TCR- $\alpha/\beta$ <sup>+</sup> CTL clone established from normal PBL. Antigen specificity of these T cell clones has not been determined.

**mAbs.** BB1 mAb (15) was generously provided by Dr. Ed Clark (University of Washington, Seattle, WA). L293 is a murine IgG1 mAb directed against CD28 that was generated by immunizing BALB/c mice with the HPB-ALL T cell line and fusing immune splenocytes with Sp2/0 myeloma cells. L303 (IgG2a,  $\kappa$ ) and L304 (IgG1,  $\kappa$ ) are murine mAbs directed against CD2 and CD2R, respectively, that were generated by immunizing BALB/c mice with anti-CD3-activated PBL and fusing immune splenocytes with Sp2/0 myeloma cells. Other mAbs were generously provided by Becton Dickinson Immunocytometry Systems (San Jose, CA). F(ab')<sub>2</sub> fragments of anti-CD3, CD28 and CD16 mAbs (all IgG1 isotype) were prepared using immobilized pepsin (Pierce Chemical Co., Rockford, IL). Purity of F(ab')<sub>2</sub> fragments was verified by SDS-PAGE analysis.

**Transfection.** P815, a murine mastocytoma cell line, and Ramos, a human Burkitt's lymphoma cell line, were obtained from American Type Culture Collection (Rockville, MD). Human B7 cDNA (generously provided by Dr. J. Allison, University of California, Berkeley) was subcloned into the pBJ expression vector (20). 10<sup>7</sup> P815 and Ramos cells were transfected with 15  $\mu$ g pBJ/B7 plasmid by electroporation (2.0 KV, 25  $\mu$ FD) using a Gene Pulser (Bio-Rad Laboratories, Richmond, CA) and selected in culture medium (RPMI 1640 [M. A. Bioproducts, Walkersville, MD] + 10% FCS [JR Scientific, Woodland, CA], 1 mM sodium pyruvate, 1 mM L-glutamine, 1% penicillin-streptomycin) supplemented with 2 mg/ml G418 (Gibco Laboratories, Grand Island, NY). After drug selection, B7-transfected P815 and Ramos cells were cloned and selected for high cell surface B7 expression by flow cytometry. Transfectants were maintained in culture medium containing 1 mg/ml G418.

**Anti-CD3-induced Redirected Cytotoxicity Assays.** Cytotoxicity

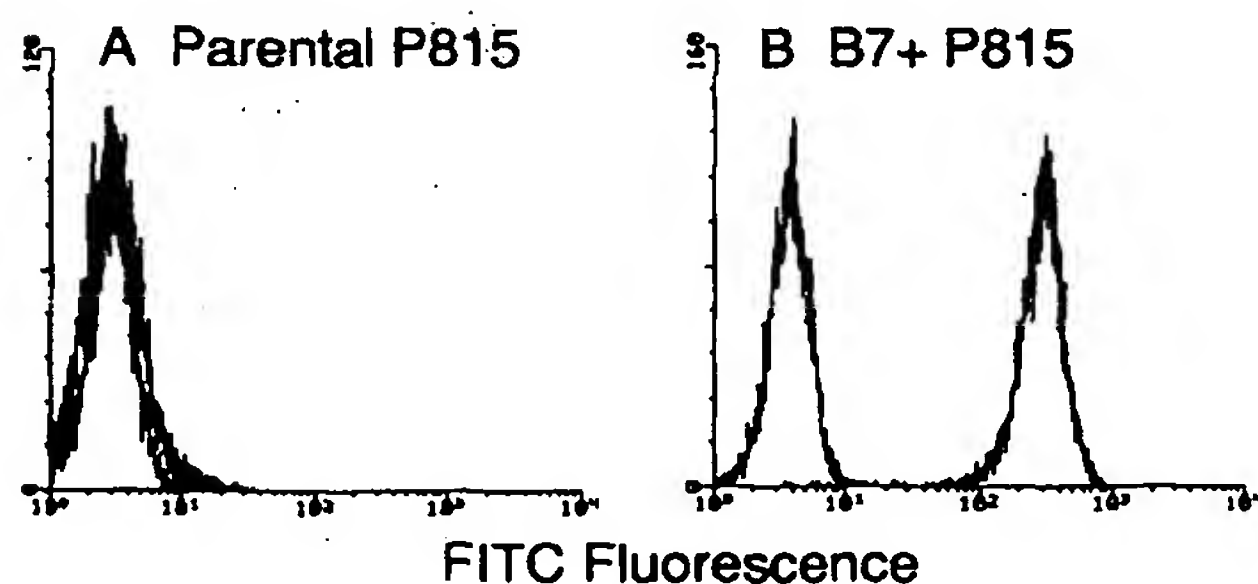
was measured in a 4-h <sup>51</sup>Cr-release assay (18). Anti-CD3 or anti-CD2 mAbs were added to <sup>51</sup>Cr-labeled P815 or B7<sup>+</sup> P815 target cells 30 min before addition of effector T cells, unless indicated otherwise. In mAb blocking studies, all mAbs were used at a final concentration of 5  $\mu$ g/ml and were added at initiation of the assay.

**Proliferation Assays.** Irradiated (80 Gy) P815 and B7<sup>+</sup> P815 ( $3 \times 10^4$  cells/well) were plated with anti-CD3 (anti-Leu-4; 5  $\mu$ g/ml) or anti-CD2 (L303 and L304, 1.25  $\mu$ g/ml) mAbs in flat-bottomed 96-well microtiter plates (Falcon Labware, Lincoln Park, NJ). After incubation for 30 min at room temperature, high buoyant density T cells were added at  $10^5$  cells/well. Cultures were incubated at 37°C with 5% CO<sub>2</sub> in a humidified atmosphere for 72 h. For the mixed lymphocyte response,  $10^5$  high buoyant density T cells and  $10^4$  irradiated (80 Gy) Ramos or B7<sup>+</sup> Ramos cells were added to microtiter plate wells with or without mAbs and were incubated for 7 d. All cultures were labeled for the final 18 h with 1  $\mu$ Ci/well [<sup>3</sup>H]thymidine (New England Nuclear, Boston, MA) and were harvested on a 96-well plate harvester (LKB Instruments Inc., Gaithersburg, MD). Incorporated radioactivity was measured in a beta plate scintillation counter (LKB Instruments, Inc.). Each value is the mean of triplicate wells. Standard mean error of the triplicates was <10%.

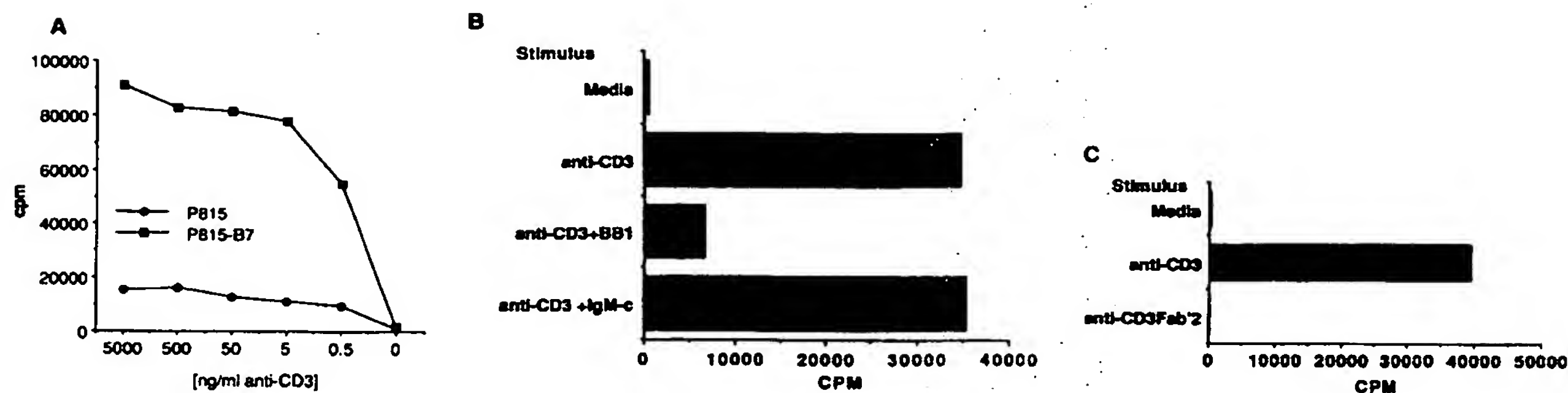
**Immunofluorescence and Flow Cytometry.** Methods of immunofluorescence and flow cytometry have been described previously (21, 22). Flow cytometry was performed using a FACScan® or FACStar<sup>PLUS</sup>® (Becton Dickinson Immunocytometry Systems).

## Results and Discussion

**CD28/B7 Costimulates Both CD2- and CD3-dependent T Cell Proliferation.** Human B7 cDNA (23) was subcloned into a mammalian expression vector, transfected into an FcR-bearing murine mastocytoma cell line, P815, and stable transfectants expressing high levels of surface B7 were established. A representative transfectant is shown in Fig. 1. To determine whether the B7<sup>+</sup> P815 cells were functionally competent to activate CD28, purified high buoyant density peripheral blood T cells were cocultured with irradiated B7<sup>+</sup> P815 or parental P815 cells in the presence or absence of anti-CD3 mAb. Con-



**Figure 1.** B7<sup>+</sup> P815 transfectants. Parental P815 (A) and B7<sup>+</sup> P815 (B) transfectants were stained with control or BB1 mAb, followed by FITC-conjugated goat anti-mouse Ig. Samples were analyzed by flow cytometry. The x-axis represents fluorescence (four-decade log scale) and the y-axis represents the relative cell number. Histograms from cells stained with control mAb (nearest the ordinate) are superimposed over the histogram of cells stained with BB1.

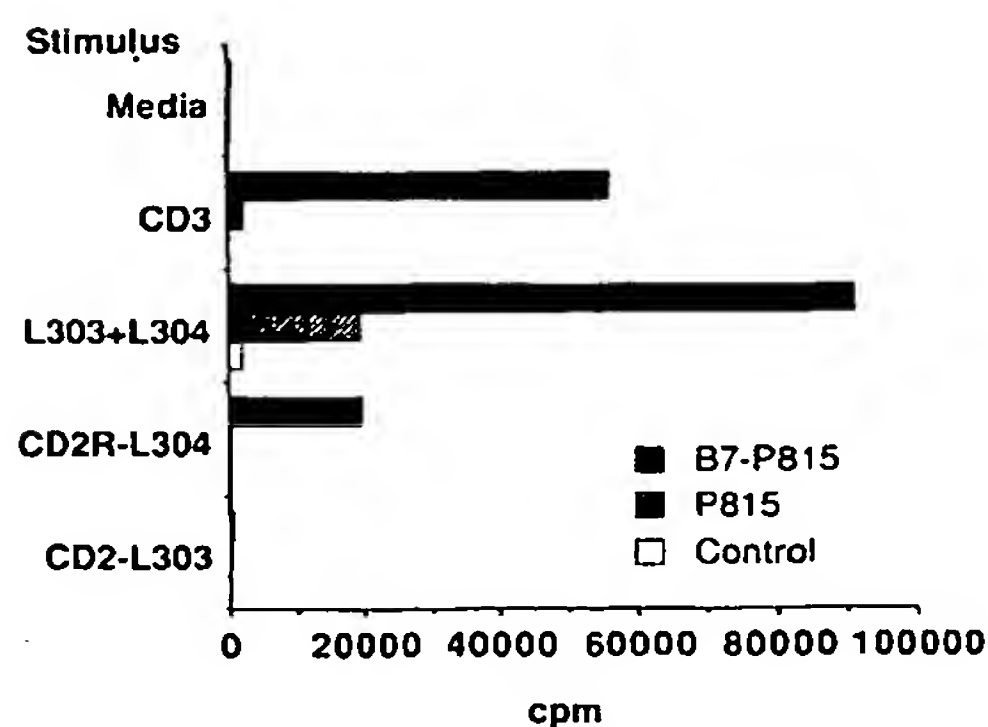


**Figure 2.** B7 costimulation of anti-CD3-induced T cell proliferation. (A) Proliferation of small, resting peripheral blood T cells cocultured with irradiated P815 (●) or B7<sup>+</sup> P815 (■) in the presence of anti-CD3 mAb. (B) BB1 mAb, but not IgM control mAb Leu-7, inhibits B7<sup>+</sup> P815 augmentation of anti-CD3-induced T cell proliferation. Note that BB1 mAb is a murine IgM and does not interfere with binding of anti-CD3 mAb to IgG Fc receptors on P815. (C) Intact but not F(ab')<sub>2</sub> fragments of anti-CD3 mAb stimulate proliferation of small, resting T cells in the presence of B7<sup>+</sup> P815.

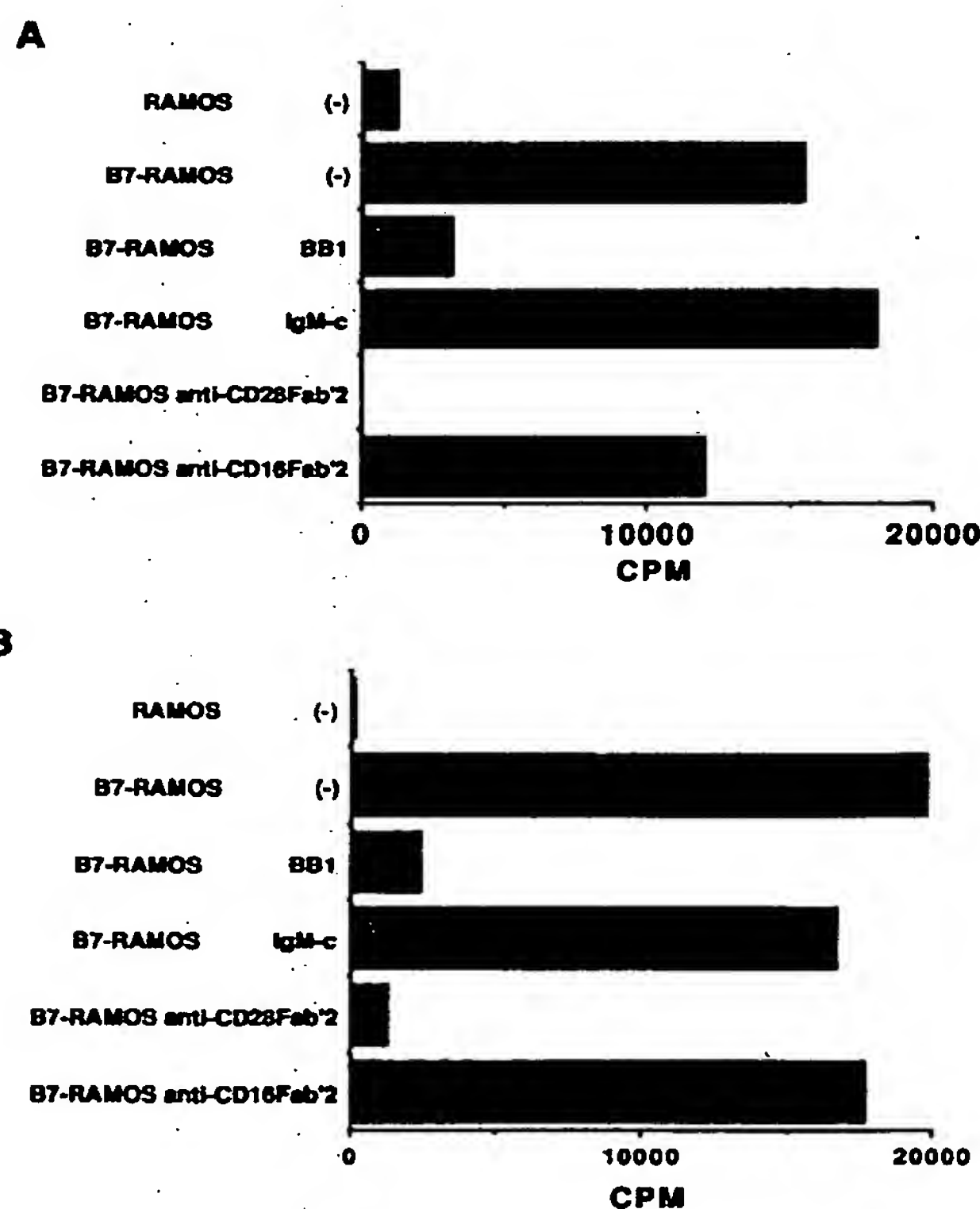
sistent with prior findings (24), freshly isolated, small peripheral blood T cells obtained from high buoyant density lymphocytes were minimally responsive to triggering via anti-CD3. However, when T cells were stimulated by coculture with B7<sup>+</sup> P815, but not parental P815, we observed a pronounced effect on anti-CD3-induced proliferation (Fig. 2 A). Anti-CD3-induced proliferation costimulated by B7<sup>+</sup> P815 was substantially inhibited by anti-B7 mAb BB1 (IgM) but not isotype-matched anti-CD57 mAb Leu-7 (IgM) (Fig. 2 B). Similarly, the addition of anti-CD28 F(ab')<sub>2</sub> fragments resulted in the inhibition of T cell proliferation (not shown). This indicated that the response was dependent upon the CD28/B7 interaction. Anti-CD3-induced proliferation was Fc dependent, in that no proliferation was observed when T cells and B7<sup>+</sup> P815 were cultured with F(ab')<sub>2</sub> anti-CD3 fragments (Fig. 2 C).

mAbs directed against certain CD2 epitopes induce T cell proliferation (25). To determine whether the CD28/B7 interaction could also stimulate T cell proliferation triggered

via CD2, we examined the effect of anti-CD2 mAbs on B7-dependent T cell proliferation (Fig. 3). Saturating concentrations of L303 (CD2) and/or L304 (CD2R) mAbs never induced substantial proliferation of small resting T cells in



**Figure 3.** B7 costimulation of anti-CD2-induced T cell proliferation. Small, resting peripheral blood T cells were cocultured with irradiated parental P815 (■), B7<sup>+</sup> P815 (■), or without P815 cells (□) in the presence or absence of anti-CD3, anti-CD2 (L303), and/or anti-CD2R (L304) mAb.



**Figure 4.** B7 expression augments an allogeneic MLR. Small, resting peripheral blood T cells from two donors (A and B) were cocultured with irradiated Ramos or B7<sup>+</sup> Ramos in the presence or absence of BB1 mAb, anti-CD28 F(ab')<sub>2</sub> fragments, control IgM mAb Leu-7, or control anti-CD16 F(ab')<sub>2</sub> fragments, as indicated.



the presence of P815. However, when resting T cells were costimulated with B7<sup>+</sup> P815, L303 and L304 together induced significant proliferation, which was inhibited by BB1 mAb (not shown). These results suggest that the binding of B7 to CD28 costimulates both the CD3 and CD2 activation pathways.

**CD28/B7 Interactions Augment Primary Allogeneic MLR.** EBV-transformed B cell lines are known to serve as potent stimulators of alloantigen-induced T cell proliferation. B lymphoblastoid cell lines have been reported to express high levels of cell surface B7 (15), suggesting the possibility that B7 expression is necessary for the generation of a primary MLR. Ramos is an EBV-negative American Burkitt's lymphoma cell line that expresses high levels of class I and II MHC antigens, but does not express B7. Preliminary experiments demonstrated that this cell line was a poor stimulator of allogeneic MLR. Ramos was transfected with B7 cDNA, and stable transfectants expressing high levels of B7 were selected. Fig. 4 shows the proliferative response of small, resting T cells from two donors cocultured with irradiated parental Ramos and B7-transfected Ramos. Parental B7-negative Ramos failed to stimulate small, resting T cells. By contrast, small, resting T cells cocultured with B7-transfected Ramos mediated substantial proliferative responses. These proliferative responses were substantially inhibited by either BB1 or anti-CD28 F(ab')<sub>2</sub> fragments, but not isotype-matched anti-CD57 or anti-CD16 F(ab')<sub>2</sub> fragments.

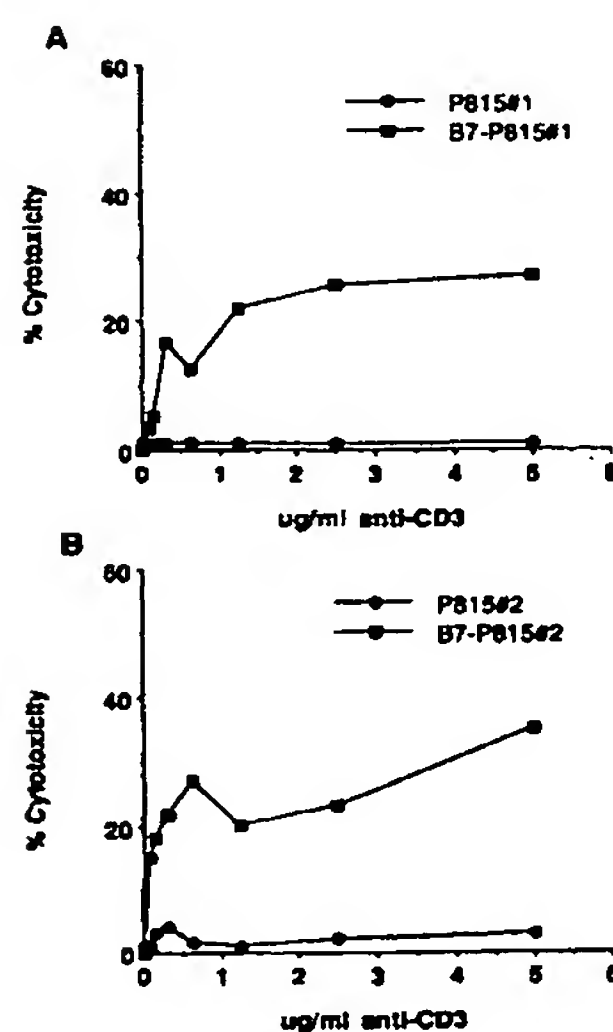
**CD28/B7 Interactions Augment Cell-mediated Cytotoxicity Triggered through CD3 and CD2.** CTL clones can lyse Fc receptor-bearing targets in the presence of anti-CD3 mAb in a "redirected" cytotoxicity assay (26–29). Unlike CTL clones, freshly isolated small resting T lymphocytes do not mediate anti-CD3 redirected lysis (29, 30), although a small subset of CD56<sup>+</sup> or CD57<sup>+</sup> T cells that may represent in vivo activated T cells do mediate this function (30, 31). Thus, en-

gagement of CD3/TCR alone is insufficient to trigger a cytolytic response in the majority of resting peripheral blood T lymphocytes. It is possible that small, resting T lymphocytes lack the cellular machinery necessary for cytotoxicity. Alternatively, additional signals may be necessary to generate a cytolytic response in resting T cells that are not required for stimulation of activated T cells. The ability of combinations of hetero-bifunctional anti-CD3/antitumor and anti-CD28/antitumor mAbs to induce cytotoxicity suggested that interaction of CD28 with its natural ligand may provide the required costimulatory signal for generation of CTL from resting peripheral blood T cells (32). As shown in Fig. 5, freshly isolated small resting peripheral blood T cells demonstrated potent anti-CD3-redirected cytotoxicity against B7<sup>+</sup> P815, but not parental P815, detectable using a 4-h <sup>51</sup>Cr release assay. Comparable results were obtained using three independently derived B7<sup>+</sup> P815 transfectants; however, P815 transfectants expressing high levels of CD16-II were not lysed in the presence of anti-CD3 (not shown). Lysis of B7<sup>+</sup> P815 transfectants was inhibited by BB1 or anti-CD28 F(ab')<sub>2</sub> fragments (Fig. 6 A). Anti-CD3-induced cytotoxicity was Fc dependent since anti-CD3 F(ab')<sub>2</sub> fragments failed to trigger lysis against B7<sup>+</sup> P815 (Fig. 6 B).

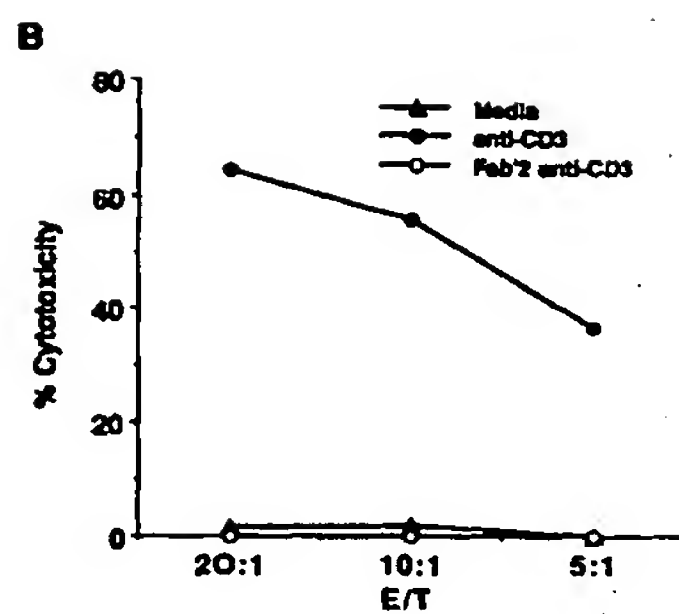
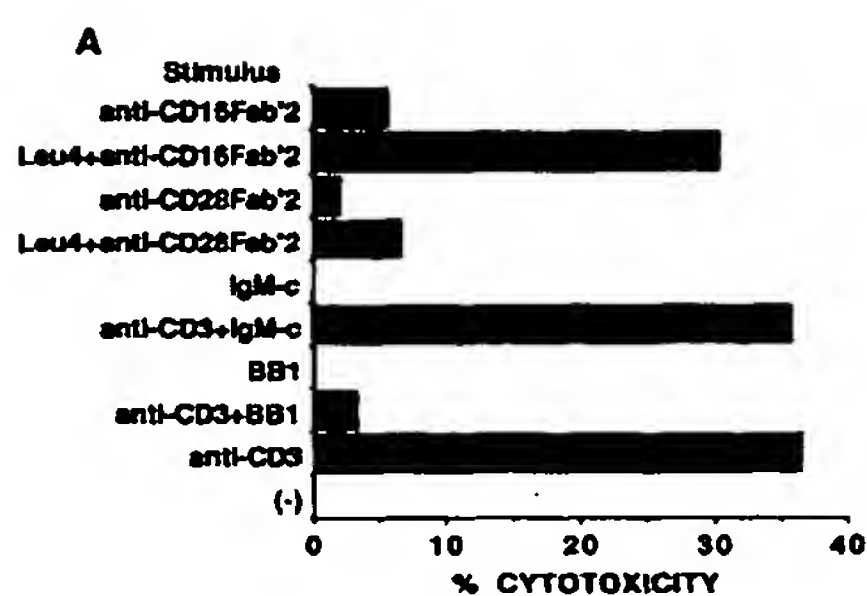
mAbs against CD4 (Leu-3a), CD5 (Leu-1), CD7 (Leu-9), CD8 (Leu-2a), CD11a (L7), CD18 (L130), and CD38 (Leu-17) failed to induce redirected lysis against parental or B7<sup>+</sup> P815 cells (not shown). However, as with the proliferative response, appropriate combinations of anti-CD2 mAb (L303 + L304) could trigger cytotoxicity against B7<sup>+</sup> P815, but not parental P815 (Fig. 7). Anti-CD2-induced cytotoxicity was inhibited by BB1 antibody, but not isotype-matched control mAb (not shown). These data indicate that binding of B7 to CD28 costimulates CD3/TCR- and CD2-dependent T cell-mediated cytotoxicity and demonstrates that small, resting T cells are capable of cytotoxic function.

The mechanism whereby expression of B7 on the target cell enables small, resting T cells to initiate lysis has not been defined. It is possible that the CD28/B7 interaction increases or stabilizes heterotypic cellular adhesion between the effector and target cells. However, in the present experimental system this seems unlikely since effectors and Fc receptor-bearing targets are already efficiently bridged by anti-CD3 or anti-CD2 mAbs. No obvious differences were observed in conjugate formation between parental or B7<sup>+</sup> P815 cells and effector T cells in the presence of anti-CD3 mAb (unpublished observation). Alternatively, interaction of CD28 with B7 on the target may provide a cosignal to the T cell that is necessary to mobilize the cellular machinery necessary for cytotoxicity. Preliminary studies have indicated that cytotoxicity is unlikely to be mediated by a soluble cytotoxic factor, since no bystander cytotoxicity was observed when <sup>51</sup>Cr-labeled parental P815 cells were admixed with unlabeled B7<sup>+</sup> P815 cells in the presence of effector T cells and anti-CD3 mAb (unpublished observation). Further studies will be necessary to elucidate the mechanism of cell-mediated cytotoxicity initiated by small, resting peripheral blood T cells.

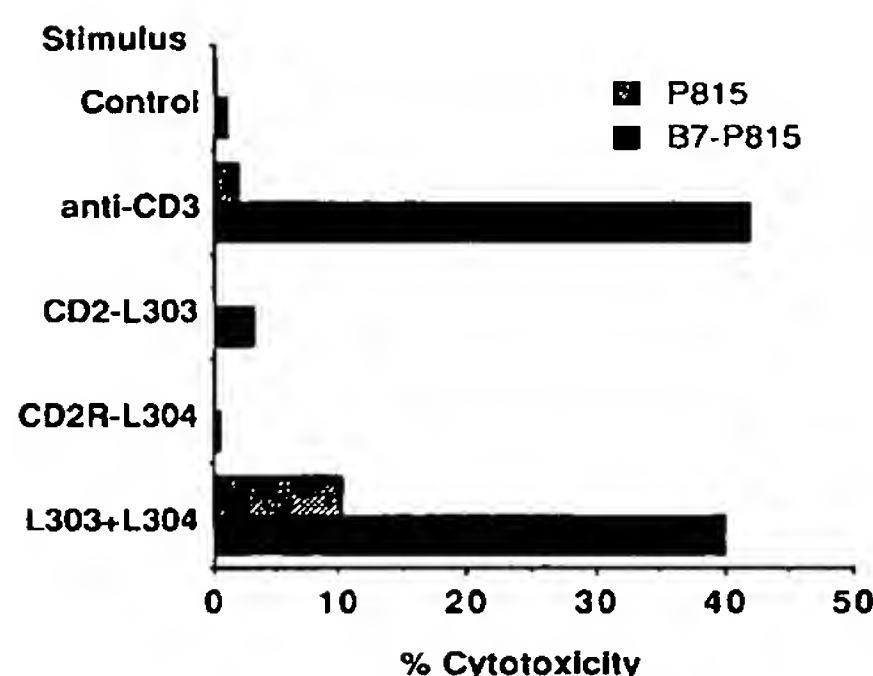
**Effectors of CD28/B7-dependent Cytotoxicity.** Using B7<sup>+</sup> P815 targets, anti-CD3-redirected cytotoxicity could also be



**Figure 5.** B7-dependent anti-CD3-redirected cytotoxicity. <sup>51</sup>Cr-labeled P815 (●) or B7<sup>+</sup> P815 (■) target cells were added to plates with titrated amount of anti-CD3 (Leu-4) mAb. After incubation for 30 min at room temperature, freshly isolated small, resting peripheral blood T cells were added and the assay was harvested after 4 h. Results from experiments using two different T cell donors (A and B) are shown. E/T ratio was 20:1.



**Figure 6.** Inhibition of B7-dependent anti-CD3-redirecited cytotoxicity. (A) Anti-CD28 F(ab')<sub>2</sub> fragments and BB1 (IgM) mAb, but not control anti-CD16 F(ab')<sub>2</sub> fragments or control Leu-7 (IgM) mAbs, inhibit anti-CD3-redirecited cytotoxicity against B7<sup>+</sup> P815. E/T ratio was 20:1. (B) Intact (●), but not F(ab')<sub>2</sub> (○) anti-CD3, induces redirecited cytotoxicity against <sup>51</sup>Cr-labeled B7<sup>+</sup> P815 targets. (▲) Cytotoxicity against B7<sup>+</sup> P815 cells in the absence of anti-CD3 mAb. All mAbs were used at a final concentration of 5 μg/ml and were added at the start of the assay.

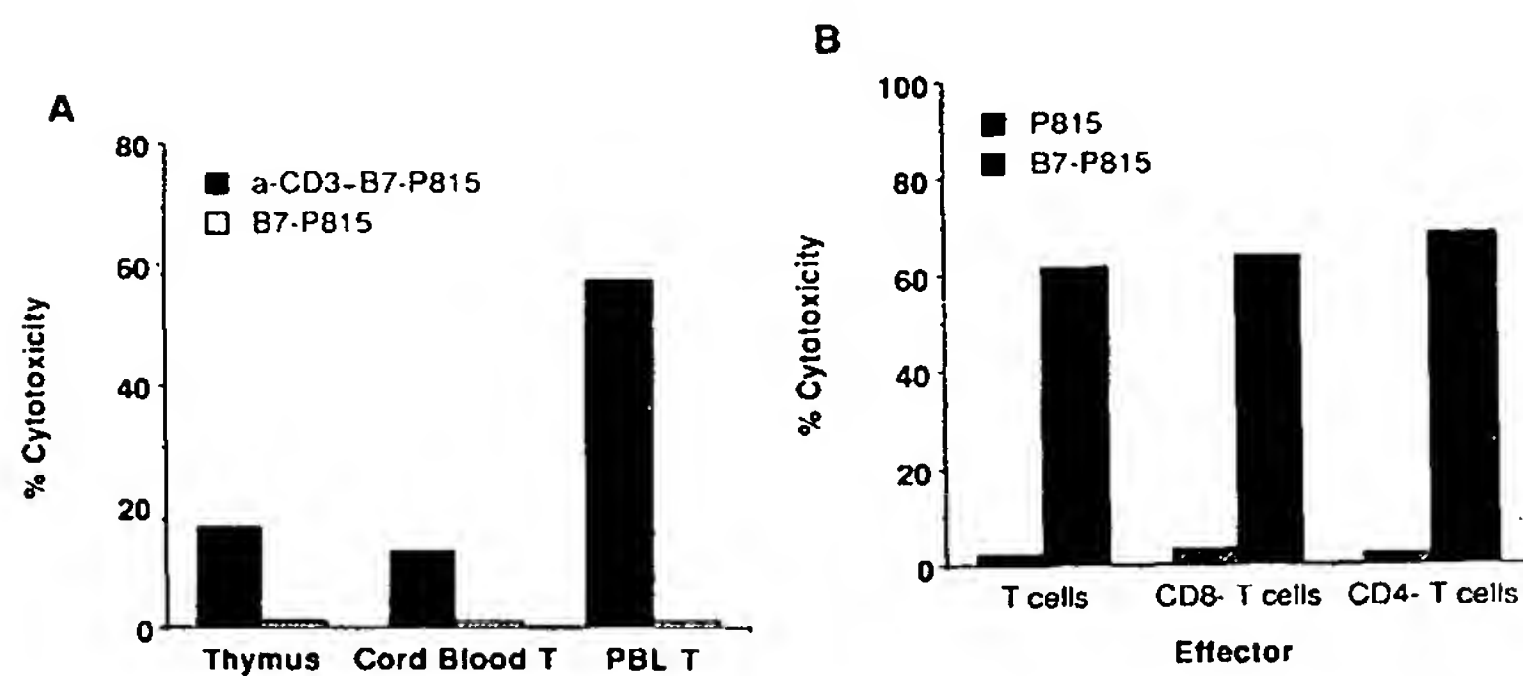


**Figure 7.** B7-dependent anti-CD2-redirecited cytotoxicity. Anti-CD2 mAb-redirecited cytotoxicity mediated by small, resting peripheral blood T cells against P815 (□) or B7<sup>+</sup> P815 (■). All mAbs were used at a final concentration of 5 μg/ml and were added at the start of the assay. E/T ratio was 20:1.

demonstrated using freshly isolated thymocyte and cord blood T cell effectors, although the activity was lower than that observed using adult small, resting T cells (Fig. 8 A). Both small resting peripheral blood CD56<sup>-</sup>, CD4<sup>+</sup> and CD56<sup>-</sup>, CD8<sup>+</sup> T cells mediated equivalent levels of anti-CD3-induced cytotoxicity (Fig. 8 B). The ability of both CD4 and

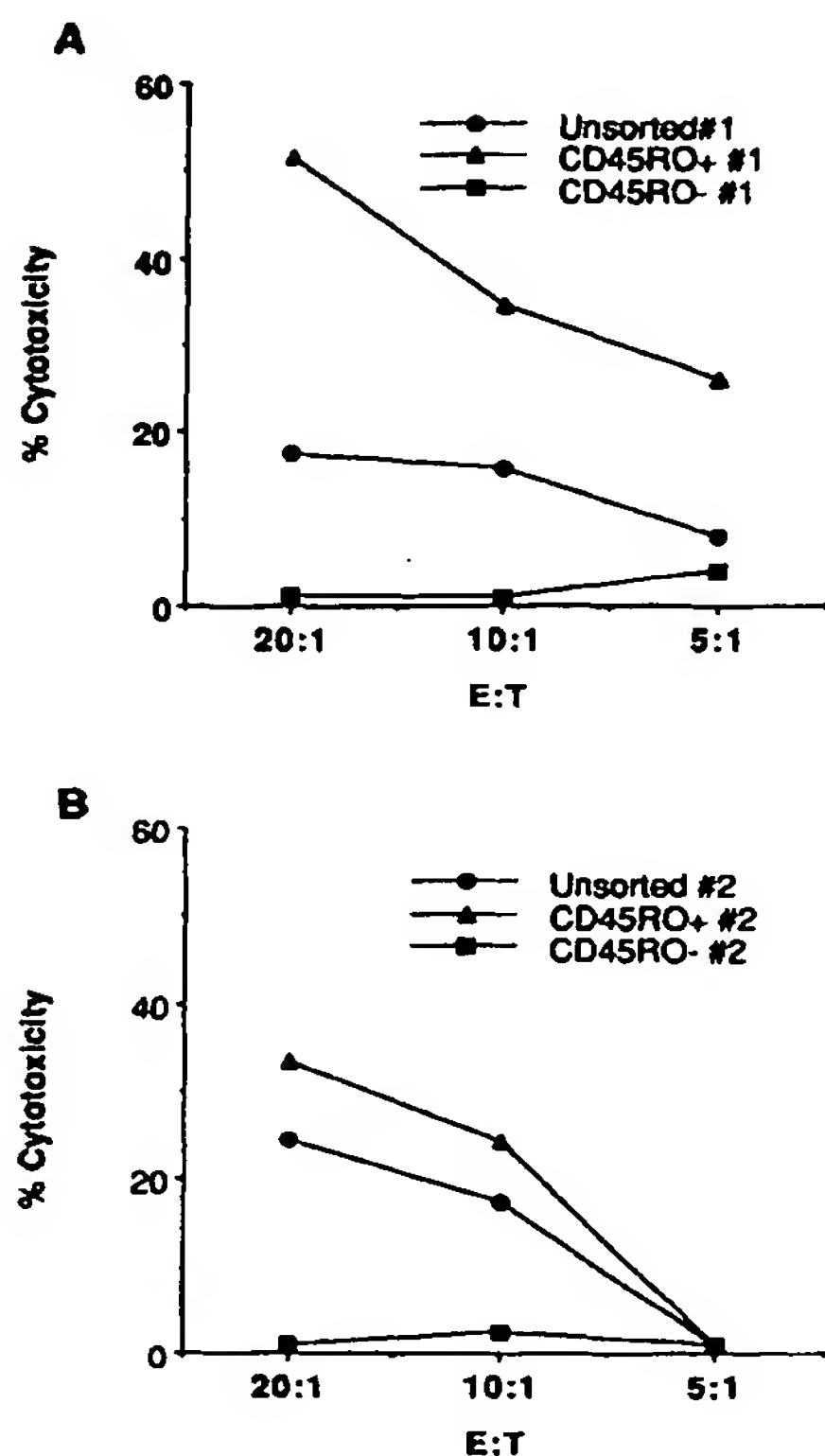
CD8 T cells to mediate cytotoxicity is consistent with the ability of both CD4 and CD8 T cell clones to mediate anti-CD3-redirecited cytotoxicity, demonstrating that both major subsets of T lymphocytes possess the necessary cellular components for cytolytic function (27, 33, 34).

Prior studies have indicated that differential immune responses can be mediated by subsets of T cells distinguished based on expression of CD45 isoforms (35, 36). "Virgin" or "naive" T cells are characterized by expression of high levels of CD45RA and the absence of CD45RO, whereas "memory" T cells express CD45RO and lower levels of CD45RA. When small, resting peripheral blood T lymphocytes were separated into CD45RO<sup>-</sup> and CD45RO<sup>+</sup> fractions, CD45RO<sup>+</sup> T cells demonstrated a significantly higher level of cytotoxicity against B7<sup>+</sup> P815 targets when stimulated with either anti-CD3 (Fig. 9) or anti-CD2 mAbs (not shown). Comparable results were obtained using T cells isolated from four different individuals. It is likely that the B7-costimulated anti-CD3-redirecited killing observed in cord blood T cells is accounted for by the presence of a low number of CD45RO<sup>+</sup> cells (9%) in this population. As shown in Fig. 10, both CD45RO<sup>-</sup> and CD45RO<sup>+</sup> T cells express CD28 on the cell surface, demonstrating that this differential responsiveness cannot be explained simply by lack of CD28 expression in the CD45RO<sup>-</sup> T cell population.



**Figure 8.** T subsets demonstrating B7-dependent T cell-mediated cytotoxicity. (A) Cytotoxicity mediated by thymocytes, cord blood T cells, and adult small, resting peripheral blood T cells against B7<sup>+</sup> P815 in the presence (■) or absence (□) of anti-CD3. Note that high buoyant density cord blood CD3<sup>+</sup> T lymphocytes were 87% CD4<sup>+</sup>, 12% CD8<sup>+</sup>, and 9% CD45RO<sup>+</sup>. T cells failed to demonstrate any cytotoxicity against parental P815 in the presence of anti-CD3 mAb (not shown). (B) Anti-CD3-redirecited cytotoxicity mediated by CD4<sup>+</sup> and CD8<sup>+</sup> small, resting peripheral blood T cells against P815 (□) or B7<sup>+</sup> P815 (■). High buoyant density T lymphocytes were stained with PE-conjugated anti-Leu-19 (CD56) and FITC-conjugated anti-Leu-2 (CD8) or anti-Leu-3 (CD4) mAbs. CD8<sup>+</sup>, CD56<sup>-</sup>; CD8<sup>+</sup>, CD56<sup>+</sup>; CD4<sup>+</sup>, CD56<sup>-</sup>, and CD4<sup>+</sup>, CD56<sup>+</sup> subsets were isolated

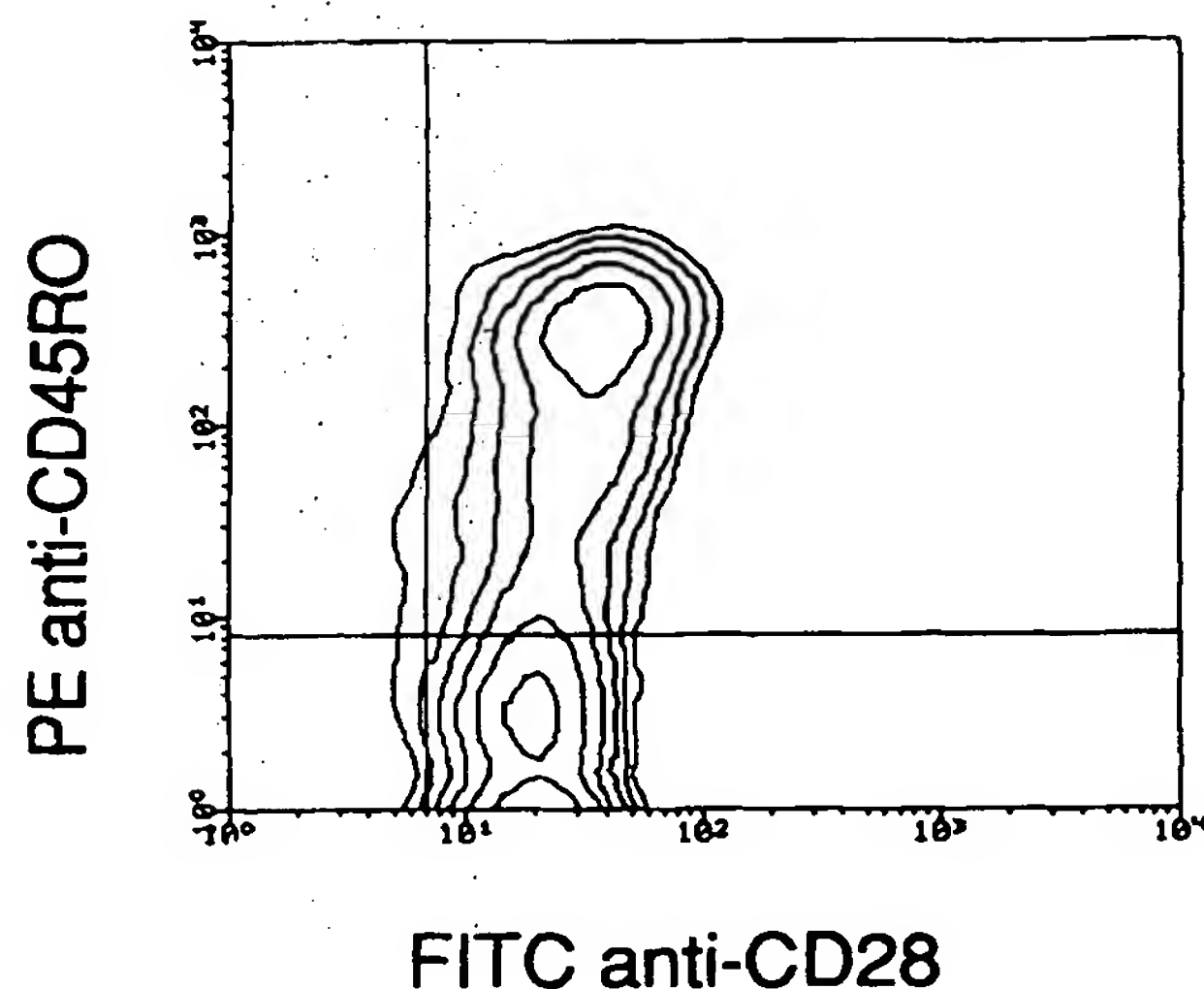
to >95% purity by flow cytometry. Data are presented for the CD4<sup>+</sup>, CD56<sup>-</sup> (i.e., CD8<sup>+</sup>) T cell fraction. Positively sorted CD8<sup>+</sup>, CD56<sup>-</sup> and CD4<sup>+</sup>, CD56<sup>-</sup> T cells were also tested and demonstrated equivalent cytotoxicity against B7-P815 targets in the presence of anti-CD3 (not shown). E/T ratio was 20:1.



**Figure 9.** B7-dependent cytotoxicity mediated by memory T cells. High buoyant density T lymphocytes from two individuals (A and B) were stained with PE-conjugated anti-CD45RO (UCHL-1) and were sorted by flow cytometry. Purity of the sorted subsets was >98%. Unseparated T cells (●), CD45RO<sup>-</sup> T cells (■), and CD45RO<sup>+</sup> T cells (▲) were assayed for anti-CD3-redirecited cytotoxicity against <sup>51</sup>Cr-labeled B7<sup>+</sup> P815 cells. Staining with anti-CD45RO mAb did not affect levels of anti-CD3-redirecited lysis compared with unstained cells (not shown).

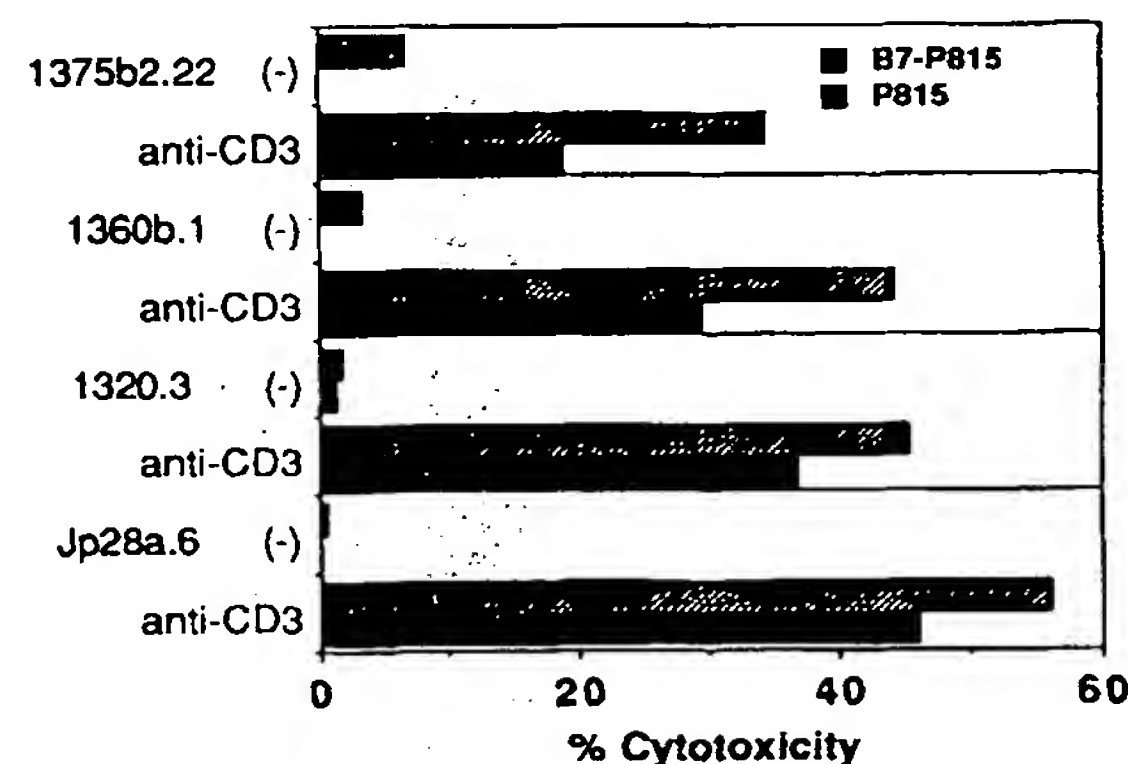
**B7-independent Cytotoxicity Mediated by T Cell Clones.** The above results suggest that small, resting T cells require CD28/B7 costimulation to induce CD2- or CD3-dependent cell-mediated cytotoxicity. In accordance with prior observations (26–29), TCR- $\alpha/\beta$ <sup>+</sup> and TCR- $\gamma/\delta$ <sup>+</sup> CTL clones mediated anti-CD3-redirecited cytotoxicity against parental P815 (Fig. 11). Thus, while resting T cells apparently require CD28/B7 interactions to trigger cell-mediated cytotoxicity, this is no longer essential for stimulation of activated T lymphocytes. However, even with the T cell clones, we consistently observed that the presence of B7 on P815 slightly augmented cytotoxicity compared with parental P815 targets (Fig. 11).

In summary, interaction of CD28 with its natural ligand B7 costimulates the induction of cell-mediated cytotoxicity, as well as proliferation, of small, resting T lymphocytes. Both CD2 and CD3 activation pathways cooperate with CD28 as a consequence of CD28 binding its natural ligand B7. With respect to induction of cytotoxicity, the CD45RO<sup>+</sup>



**Figure 10.** Coexpression of CD28 on CD45RO<sup>-</sup> and CD45RO<sup>+</sup> small, resting peripheral blood T cells. High buoyant density T lymphocytes were stained with FITC-conjugated anti-CD28 (9.3), PE-conjugated anti-CD45RO, and PerCP-conjugated anti-CD3, or appropriate control mAb. Samples were analyzed on a FACScan<sup>®</sup>. Markers were positioned to include >98% of cells stained with FITC- and PE-conjugated isotype control antibodies in the lower left quadrant (not shown). An electronic gate was set on PerCP-stained CD3<sup>+</sup> T cells, and correlated fluorescence of FITC (CD28) and PE (CD45RO) were displayed as a contour map (four-decade log scales). Analysis of CD45RO expression in small, resting T cells isolated from four individuals revealed that the proportion of CD45RO<sup>+</sup> T cells ranged from 48 to 78%.

memory T cell population is primarily responsible for this activity and both CD4 and CD8 T cells are capable of this function. It is unlikely that this cytotoxicity is mediated by in vivo activated T lymphocytes, since highly purified, small,



**Figure 11.** B7-independent anti-CD3-redirecited cytotoxicity mediated by T cell clones. TCR- $\alpha/\beta$ <sup>+</sup> (Jp28a.6, 1360b.1, and 1320.3) and TCR- $\gamma/\delta$ <sup>+</sup> (1375b2.22) T cell clones were incubated with <sup>51</sup>Cr-labeled B7<sup>+</sup> P815 (■) or P815 (□) in the presence or absence of anti-CD3 mAb, as indicated. E/T ratio was 10:1.



resting T cells isolated by density gradient centrifugation were used for these experiments. Moreover, cytotoxicity was only observed using the B7<sup>+</sup> P815 transfectant and not with the parental P815 cells. Since in vitro cultured CTL, as well as the presumably in vivo activated CD56<sup>+</sup> T cells mediate redirected lysis against parental P815, it seems likely that the CD28/B7 interaction is necessary only for costimulation of small, resting T cells and is no longer required after activation. Consistent with this concept, we have previously demonstrated that TCR- $\gamma/\delta$ <sup>+</sup> T lymphocytes that mediate potent anti-CD3-redredirected cytotoxicity downregulate CD28 expres-

sion after stimulation with IL-2 (37). Moreover, the subset of CD28<sup>-</sup> peripheral blood T cells that coexpress CD11b in vivo are predominantly granular, low buoyant density T cells that may represent activated T cells that have undergone clonal expansion (38, 39). Therefore, the CD28/B7 interaction may be most critical in the early generation of a cell-mediated immune response, consistent with a possible role of CD28/B7 in the stimulation of immature thymocytes (40, 41). The existence of functional cytolytic T cells in the small, resting T cell population may be advantageous in facilitating rapid responses to immune challenge.

We thank Drs. A. Jackson, E. Clark, J. Hansen, H. Spits, Y.-h. Chien, M. Davis, and J. Allison for generously providing antibodies, tissues and vectors. We also thank Dr. J. Cupp, Ms. D. Polakoff, and Ms. Jean Herrman for expert assistance with flow cytometry, and Ms. Jill Zahner for assistance with manuscript preparation.

DNAX is supported by Schering-Plough Corporation.

Address correspondence to Dr. Lewis L. Lanier, Department of Immunology, DNAX Research Institute, 901 California Avenue, Palo Alto, CA 94304

Received for publication 17 September 1991.

## References

1. Imboden, J.B., and J.D. Stobo. 1985. Transmembrane signalling by the T cell antigen receptor: perturbation of the T3-antigen receptor complex generates inositol phosphates and releases calcium ions from intracellular stores. *J. Exp. Med.* 161:446.
2. Weiss, A., J. Imboden, D. Shoback, and J. Stobo. 1984. Role of T3 surface molecule in human T-cell activation: T3-dependent activation results in an increase in cytoplasmic free calcium. *Proc. Natl. Acad. Sci. USA.* 81:4169.
3. Weiss, M.J., J.F. Daley, J.C. Hodgdon, and E.L. Reinherz. 1984. Calcium dependency of antigen-specific (T3-Ti) and alternative (T11) pathways of human T-cell activation. *Proc. Natl. Acad. Sci. USA.* 81:6836.
4. Weiss, A., J. Imboden, K. Hardy, B. Manger, C. Terhorst, and J. Stobo. 1986. The role of the T3/antigen receptor complex in T-cell activation. *Annu. Rev. Immunol.* 6:593.
5. Altmann, D.M., N. Hogg, J. Trowsdale, and D. Wilkinson. 1989. Cotransfection of ICAM-1 and HLA-DR reconstitutes human antigen-presenting cell function in mouse L cells. *Nature (Lond.)* 338:512.
6. Hansen, J.A., P.J. Martin, and R.C. Nowinski. 1980. Monoclonal antibodies identifying a novel T-cell antigen and Ia antigens of human lymphocytes. *Immunogenetics.* 10:247.
7. Hara, T., S.M. Fu, and J.A. Hansen. 1985. Human T cell activation II. A new activation pathway used by a major T cell population via a disulfide-bonded dimer of a 44 kilodalton polypeptide (9.3 antigen). *J. Exp. Med.* 161:1513.
8. Martin, P.J., J.A. Ledbetter, Y. Morishita, C.H. June, P.G. Beatty, and J.A. Hansen. 1986. A 44 kilodalton cell surface homodimer regulates interleukin 2 production by activated human T lymphocytes. *J. Immunol.* 136:3282.
9. Yang, S.Y., S.M. Denning, S. Mizuno, B. Dupont, and B.F. Haynes. 1988. A novel activation pathway for mature thymocytes: costimulation of CD2 (Tp50) and CD28 (Tp44) induces autocrine interleukin 2/interleukin 2 receptor-mediated cell proliferation. *J. Exp. Med.* 168:1457.
10. Damle, N.K., L.V. Doyle, L.S. Grosmaire, and J.A. Ledbetter. 1988. Differential regulatory signals delivered by antibody binding to the CD28 (Tp44) molecule during the activation of human T lymphocytes. *J. Immunol.* 140:1753.
11. van Lier, R.A., M. Brouwer, and L.A. Aarden. 1988. Signals involved in T cell activation. T cell proliferation induced through the synergistic action of anti-CD28 and anti-CD2 monoclonal antibodies. *Eur. J. Immunol.* 18:167.
12. Pierres, A., M. Lopez, C. Cerdan, J. Nunes, D. Olive, and C. Mawas. 1988. Triggering CD28 molecules synergize with CD2 (T11.1 and T11.2)-mediated T cell activation. *Eur. J. Immunol.* 18:685.
13. Lindsten, T., C.H. June, J.A. Ledbetter, G. Stella, and C.B. Thompson. 1989. Regulation of lymphokine messenger RNA stability by a surface-mediated T cell activation pathway. *Science (Wash. DC).* 244:339.
14. Freedman, A.S., G. Freeman, J.C. Horowitz, J. Daley, and L.M. Nadler. 1987. B7, a B cell-restricted antigen that identifies preactivated B cells. *J. Immunol.* 139:3260.
15. Yokochi, T., R.D. Holly, and E.A. Clark. 1982. B lymphoblast antigen (BB-1) expressed on Epstein-Barr virus-activated B cell blasts, B lymphoblastoid cell lines, and Burkitt's lymphomas. *J. Immunol.* 128:823.
16. Linsley, P.S., W. Brady, L. Grosmaire, A. Aruffo, N.K. Damle, and J.A. Ledbetter. 1991. Binding of the B cell activation antigen B7 to CD28 costimulates T cell proliferation and interleukin 2 mRNA accumulation. *J. Exp. Med.* 173:721.

17. Koulouva, L., E.A. Clark, G. Shu, and B. Dupont. 1991. The CD28 ligand B7/BB1 provides costimulatory signal for alloactivation of CD4<sup>+</sup> T cells. *J. Exp. Med.* 173:759.
18. Phillips, J.H., and L.L. Lanier. 1986. Dissection of the lymphokine-activated killer phenomenon: relative contribution of peripheral blood natural killer cells and T lymphocytes to cytotoxicity. *J. Exp. Med.* 164:814.
19. Yssel, H., J.E. De Vries, M. Koken, W. van Blitterswijk, and H. Spits. 1984. Serum-free medium for the generation and the propagation of functional human cytotoxic and helper T cell clones. *J. Immunol. Methods.* 72:219.
20. Lin, A.Y., B. Devaux, A. Green, C. Sagerstrom, J.F. Elliott, and M.M. Davis. 1990. Expression of T cell antigen receptor heterodimers in a lipid-linked form. *Science (Wash. DC)* 249:677.
21. Lanier, L.L., A.M. Le, J.H. Phillips, N.L. Warner, and G.F. Babcock. 1983. Subpopulations of human natural killer cells defined by expression of the Leu-7 (HNK-1) and Leu-11 (NK-15) antigens. *J. Immunol.* 131:1789.
22. Lanier, L.L., and D.J. Recktenwald. 1991. Multicolor immunofluorescence and flow cytometry. *Methods, A Companion to Methods Enzymol.* 2:192.
23. Freeman, G.J., A.S. Freedman, J.M. Segil, G. Lee, J.F. Whitman, and L.M. Nadler. 1989. B7, a new member of the Ig superfamily with unique expression on activated and neoplastic B cells. *J. Immunol.* 143:2714.
24. Ledbetter, J.A., C.H. June, P.J. Martin, C.E. Spooner, J.A. Hansen, and K.E. Meier. 1986. Valency of CD3 binding and internalization of the CD3 cell-surface complex control T cell responses to second signals: distinction between effects of protein kinase C, cytoplasmic free calcium, and proliferation. *J. Immunol.* 136:3945.
25. Meuer, S.C., R.E. Hussey, M. Fabbi, D. Fox, O. Acuto, K.A. Fitzgerald, J.C. Hodgdon, J.P. Protentis, S.F. Schlossman, and E.L. Reinherz. 1984. An alternative pathway of T-cell activation: a functional role for the 50 kD T11 sheep erythrocyte receptor protein. *Cell.* 36:897.
26. Spits, H., H. Yssel, J. Leeuwenberg, and J.E. DeVries. 1985. Antigen-specific cytotoxic T cell and antigen-specific proliferating T cell clones can be induced to cytolytic activity by monoclonal antibodies against T3. *Eur. J. Immunol.* 15:88.
27. Mentzer, S.J., J.A. Barbosa, and S.J. Burakoff. 1985. T3 monoclonal antibody activation of nonspecific cytotoxicity: a mechanism of CTL inhibition. *J. Immunol.* 135:34.
28. Siliciano, R.F., J.C. Pratt, R.E. Schmidt, J. Ritz, and E.L. Reinherz. 1985. Activation of cytolytic T lymphocyte and natural killer cell function through the T11 sheep erythrocyte binding protein. *Nature (Lond.)* 317:428.
29. Leeuwenberg, J.F.M., H. Spits, W.J.M. Tax, and J.A. Capel. 1985. Induction of nonspecific cytotoxicity by monoclonal anti-T3 antibodies. *J. Immunol.* 134:3770.
30. Phillips, J.H., and L.L. Lanier. 1986. Lectin-dependent and anti-CD3 induced cytotoxicity are preferentially mediated by peripheral blood cytotoxic T lymphocytes expressing Leu-7 antigen. *J. Immunol.* 136:1579.
31. Garrido, M.A., P. Perez, J.A. Titus, M.J. Valdayo, D.F. Winkler, S.A. Barbieri, J.R. Wunderlich, and D.M. Segal. 1990. Targeted cytotoxic cells in human peripheral blood lymphocytes. *J. Immunol.* 144:2891.
32. Jung, G., J.A. Ledbetter, and H.J. Muller-Eberhard. 1987. Induction of cytotoxicity in resting human T lymphocytes bound to tumor cells by antibody heteroconjugates. *Proc. Natl. Acad. Sci. USA.* 84:4611.
33. Hayward, A., A. Boylston, and P. Beverley. 1988. Lysis of CD3 hybridoma targets by cloned human CD4 lymphocytes. *Immunology.* 64:87.
34. Rottevel, F.T.M., I. Kokkelink, R.A.W. van Lier, B. Kuenen, A. Meager, F. Miedema, and C.J. Lucas. 1988. Clonal analysis of functionally distinct human CD4<sup>+</sup> T cell subsets. *J. Exp. Med.* 168:1659.
35. Sanders, M.E., M.W. Makgoba, and S. Shaw. 1988. Human naive and memory T cells: reinterpretation of helper-inducer and suppressor-inducer subsets. *Immunol. Today.* 9:195.
36. Beverley, P.C.L. 1990. Is T-cell memory maintained by cross-reactive stimulation? *Immunol. Today.* 11:203.
37. Testi, R., and L.L. Lanier. 1989. Functional expression of CD28 on T cell antigen receptor  $\gamma/\delta$ -bearing T lymphocytes. *Eur. J. Immunol.* 19:185.
38. Morishita, Y., H. Sao, J.A. Hansen, and P.J. Martin. 1989. A distinct subset of human CD4<sup>+</sup> cells with a limited alloreactive T cell receptor repertoire. *J. Immunol.* 143:2783.
39. Yamada, H., P.J. Martin, M.P. Braun, P.G. Beatty, K. Sadamoto, and J.A. Hansen. 1985. Monoclonal antibody 9.3 and anti-CD11 antibodies define reciprocal subsets of lymphocytes. *Eur. J. Immunol.* 15:1164.
40. Turka, L.A., J.A. Ledbetter, K. Lee, C.H. June, and C.B. Thompson. 1990. CD28 is an inducible T cell surface antigen that transduces a proliferative signal in CD3<sup>+</sup> mature thymocytes. *J. Immunol.* 144:1646.
41. Turka, L.A., P.S. Linsley, R. Paine III, G.L. Schieven, C.B. Thompson, and J.A. Ledbetter. 1991. Signal transduction via CD4, CD8, and CD28 in mature and immature thymocytes: Implications for thymic selection. *J. Immunol.* 146:1428.

## Two Distinct Intracytoplasmic Regions of the T-cell Adhesion Molecule CD28 Participate in Phosphatidylinositol 3-Kinase Association\*

(Received for publication, December 1, 1995, and in revised form, February 1, 1996)

Françoise Pagès†§, Marguerite Ragueneau‡, Sandrine Klasen‡, Michela Battifora†¶, Dominique Couez‡||, Ray Sweet\*\*, Alemseged Truneh\*\*, Stephen G. Ward‡ §§, and Daniel Olive†§§

From †INSERM Unit 119, 27 bd Leï Roure, 13009 Marseille, France and \*\*SmithKline Beecham Pharmaceuticals, King of Prussia, Pennsylvania 19406

Through the interaction with its ligands, CD80/B7-1 and CD86/B7-2 or B70, the human CD28 molecule plays a major functional role as a costimulator of T cells along with the CD3-TcR complex. We and others have previously reported that phosphatidylinositol 3-kinase inducibly associates with CD28. This association is mediated by the SH2 domains of the p85 adaptor subunit interacting with a cytoplasmic YNM consensus motif present in CD28 at position 173-176. Disruption of this binding site by site-directed mutagenesis abolishes CD28-induced activation events in a murine T-cell hybridoma transfected with human CD28 gene.

Here we show that the last 10 residues of the intracytoplasmic domain of CD28 (residues 193-202) are required for its costimulatory function. These residues are involved in interleukin-2 secretion, p85 binding, and CD28-associated phosphatidylinositol 3-kinase activity. In contrast, the CD28/CD80 interaction is unaffected by this deletion, as is the induction of other second messengers such as the rise in intracellular calcium and tyrosine phosphorylation of CD28-specific substrates. Furthermore, we also demonstrate that, within these residues, the tyrosine at position 200 is involved in p85 binding, probably together with the short proline-rich motif present between residues 190 and 194 (PYAPP).

In the absence of a costimulatory signal, activation of the CD3-TcR<sup>1</sup> complex is not sufficient to induce the complete activation of T lymphocytes. The interaction between CD28 on T lymphocytes and its counter-receptors CD80 (B7-1) and

CD86 (B70 or B7-2) on antigen-presenting cells provides a costimulatory signal required for IL-2 production, T-cell proliferation, and effector functions such as T-cell-mediated cytotoxicity and differentiation of Th cells into Th1 or Th2 subsets (for recent reviews, see Refs. 1-3). This CD28/CD80 interaction has also been shown to prevent anergy and to boost anti-tumor immunity (4-6).

Sequence comparisons between human, rat, mouse, and chicken CD28 cytoplasmic domains (7-10) demonstrates high interspecies conservation, suggesting a crucial role for this domain in coupling to signal transduction pathways. In the absence of catalytic motifs in this sequence, an indirect coupling via adaptor molecules was the most likely mechanism of action. Indeed, we and others have demonstrated previously that ligand stimulation of the human CD28 molecule induces its association with PI 3-kinase activity (11-15) by means of a cytoplasmic YNM motif at position 173-176 which, when phosphorylated, interacts with the SH2 domains of the p85 adaptor subunit. Similarly, the SH2 domain of the adaptor protein Grb-2 has been shown to interact with this motif although with a lower affinity (16), and the CD28-associated Grb-2-Sos complexes are likely to link the activated CD28 receptor to the activation of p21<sup>ras</sup> and downstream events such as Raf-1 hyperphosphorylation and ERK2 stimulation (17), as well as Jun kinase activation (18).

The primary events leading to CD28 phosphorylation are becoming better understood. The T-cell-specific protein-tyrosine kinase ITK has been shown to associate with CD28 and to be phosphorylated on tyrosine residues after CD28 stimulation (19), and the Src-related tyrosine kinases p56<sup>lck</sup> and p59<sup>lyn</sup> have been found in CD28 immune complexes from stimulated T cells (20). Recently, it has been shown that CD28 is phosphorylated by p56<sup>lck</sup> and p59<sup>lyn</sup> *in vitro* leading to the recruitment of ITK, Grb-2, and p85 (21). Interestingly, the pattern of tyrosine phosphoproteins induced by a CD28 stimulation is similar but not superimposable to that induced by a CD3-TcR stimulation (22, 23) and, among the identified products, are p36-38, p95<sup>vav</sup>, and PLC-γ1 as well as a CD28-specific 64-kDa protein which has yet to be formally identified (reviewed in Ref. 24).

Using a murine T-cell hybridoma transfected with the human CD28 gene, we have shown previously that a point mutation of the Tyr<sup>173</sup> residue into phenylalanine abolished CD28-induced IL-2 secretion, suggesting that the PI 3-kinase pathway plays a major role in the CD28 function (12). Here we report the generation and functional characterization of a set of intracytoplasmic variants of the human CD28 molecule. We have generated mutants of CD28 containing progressive truncations of its intracytoplasmic tail (10, 21, 30, and 41 residues), as well as a point mutation of the tyrosine residue at position 200. These variants were expressed in a murine T-cell hybridoma

\* This work was supported in part by Grant ERB CHRX CT94-0537 from the EC and grants from the Association pour la Recherche contre le Cancer and Ligue Nationale contre le Cancer. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Present address: Ludwig Institute for Cancer Research, University College of London, 91 Riding House St., London W1 8BT, UK.

¶ Present address: DIMI Università di Genova, viale Benedetto, 16132 Genova, Italy.

|| Present address: INSERM Unit 298, CHRU-F49033, Angers Cedex 01, France.

§§ Recipient of an INSERM fellowship (poste vert). Present address: School of Pharmacy and Pharmacology, University of Bath, Claverton Down, Bath, Avon BA2 7AY, UK.

§§§ To whom correspondence should be addressed. Tel.: 33-9175-8415; Fax: 33-9126-0364.

<sup>1</sup> The abbreviations used are: TcR, T-cell receptor; PI 3-kinase, phosphatidylinositol 3-kinase; IL-2, interleukin 2; Th, T helper; SH, Src homology; Ig, immunoglobulin; PLC, phospholipase C; mAb, monoclonal antibody; PAGE, polyacrylamide gel electrophoresis; GST, glutathione S-transferase; HPLC, high performance liquid chromatography.



doma. By analyzing stable transfectants, we investigated whether these molecules were able to mediate cell adhesion to human CD80-transfected L-cells, to be phosphorylated, bind and activate PI 3-kinase, and to costimulate IL-2 production together with CD3·TcR.

#### EXPERIMENTAL PROCEDURES

**Cells and mAbs**—DC27.1 used for transfection is a murine T-cell hybridoma derived by transfecting the TcR  $\alpha\beta$  genes of KB<sub>5</sub>C<sub>20</sub> in DO11.10.2 (kindly provided by B. Malissen, CIML, France). These cells were grown in Dulbecco's modified Eagle's medium containing 10% fetal calf serum, sodium pyruvate (1 mM),  $\beta$ -mercaptoethanol (50  $\mu$ M), and antibiotics (penicillin-streptomycin, 10 IU/ml), supplemented by xanthine (250  $\mu$ g/ml), hypoxanthine (13.6  $\mu$ g/ml), and mycophenolic acid (2  $\mu$ g/ml). LTK<sup>−</sup> and LB7<sup>+</sup> cells are L cells respectively untransfected or transfected by a CD80/B7-containing expression vector.<sup>2</sup> The human CD28 mAbs, CD28.1, CD28.2, CD28.3, CD28.5, CD28.6, and 248 used in this study have been described previously (25). C11E.4 and 6A11.2 (anti-human IgG1 and IgM, respectively) were derived in the laboratory and used as negative isotypic controls. The anti-murine mAbs were, respectively, 145-2C11 (a hamster IgG specific for CD3- $\epsilon$  chain) and 37.51 (specific for murine CD28, Pharmingen, San Diego, CA).

**Oligonucleotide-directed Mutagenesis**—The human CD28 cDNA (kind gift of B. Seed, Ref. 7) was cloned into the *Sa*II-*Bam*HI open vector pBluescript KS<sup>+</sup>. We constructed deletion mutants using the "Mutagenesis Phagemid *In Vitro* Mutagenesis Kit" (Bio-Rad). The point mutation of the Tyr<sup>200</sup> residue to phenylalanine was realized with the overlap extension technique. The double mutant Tyr<sup>173-200</sup> was obtained by using these primers on the cDNA template containing the mutation Tyr<sup>173</sup> → Phe<sup>173</sup> (12). The sequences of the oligonucleotides used are available upon request. Sequencing of mutated molecules was performed using Sequenase 2.0 (U. S. Biochemical Corp.).

**Plasmid Construction and Transfection**—Wild type and mutated CD28 cDNA constructs were cloned into pHAPr-1-neo (26) at *Sa*II/*Bam*HI sites, and recombinant genes were introduced by protoplast fusion into DC27.1 as described (27). Stable transfectants were selected for their resistance to 3 mg/ml geneticin G418 (Life Technologies, Inc.) and screened for CD28 expression by flow cytometry analysis.

**Flow Cytometry Analysis**— $2 \times 10^5$  cells were incubated with saturating concentrations of mAbs at 4 °C for 1 h. After extensive washing, cells were stained with fluorescein isothiocyanate-conjugated goat anti-mouse Ig at 4 °C for 30 min (Jackson Laboratories, West Grove, PA). Samples were analyzed by flow cytometry using a FACScan (Becton Dickinson). Fluorescence data were collected with logarithmic amplification.

**Adhesion Assay**— $4 \times 10^5$  transfected cells loaded with calcein AM (Molecular Probes, Eugene, OR) were added to  $10^5$  LTK<sup>−</sup> or LB7<sup>+</sup> cells seeded the day before in a microtiter plate, in the absence or presence of mAb CD28.2 in PBS without Ca<sup>2+</sup> and Mg<sup>2+</sup>. Adherent cells were analyzed by the quantification of fluorescence (excitation at 485 nm and emission at 538 nm) by fluorimetry (Fluoroscanner).

**Measurement of IL-2 Secretion**— $10^5$  transfected cells were stimulated for 24 h at 37 °C in microtiter plates with various stimuli.  $5 \times 10^4$  untransfected (LTK<sup>−</sup>) or CD80-transfected (LB7<sup>+</sup>) L cells were used for stimulation of transfectants. Negative (anti-CD5) and positive (anti-CD3) controls were respectively purified mAbs C11E.4 and 145-2C11 used at a final concentration of 10  $\mu$ g/ml in combination with FcR<sup>+</sup> B lymphoma cells LK35.2 ( $10^5$  cells). Soluble CD28.2 mAb (30  $\mu$ g/ml) was used in combination with soluble 145-2C11 (10  $\mu$ g/ml). Culture supernatants were collected and titrated, by serial 2-fold dilutions, for their ability to support proliferation of the IL-2-dependent murine T cell line, CTLL-2, as assessed by the cell growth determination 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide kit (Sigma). Results were expressed as  $A_{570\text{ nm}}$  obtained for each dilution of the supernatants.

**PI 3-Kinase Assay**— $10^7$  transfected cells were either unstimulated or activated by CD28.2, then immunoprecipitated with protein G-Sepharose beads (Pharmacia Biotech Inc.), or with a p85 antiserum (UBI). Measure of CD28-associated PI 3-kinase activity was performed as described in Ref. 12.

**Association of p85 Subunit and CD28 Phosphorylation**— $10^7$  cells were stimulated by a fibroblast cell line transfected (LB7<sup>+</sup>) or not (LTK<sup>−</sup>) by the CD80/B7 cDNA, with or without CTLA-4Ig. Western blotting of p85-associated CD28 molecules was performed as described

in Ref. 12. Blots were then stripped and reprobed with an anti-phosphotyrosine antibodies (4G10, UBI). Integrated signal intensity of phosphorylated proteins was determined using the BioImage System (Mili-pore) and expressed as arbitrary units. Values were normalized by integration of the actual amount of CD28 molecule detected by Western blotting for each mutant. Fold induction corresponds to the ratio of CD28 tyrosine phosphorylation in cells stimulated by CD80<sup>+</sup>-L cells versus unstimulated cells.

**Anti-phosphotyrosine Immunoblotting**—Whole cell lysates from  $2 \times 10^6$  stimulated cells were run on SDS-PAGE and transferred to polyvinylidene difluoride membranes as described in Ref. 12. Blots were then probed with an anti-phosphotyrosine antibodies (4G10, UBI), and immunoreactive proteins were visualized by ECL.

**Binding of CD28 Peptides to Purified p85 Domains**—CD28 peptides (residues 166–180, 166–180 with phosphorylated Tyr<sup>173</sup>, 186–202, 186–202 with phosphorylated Tyr<sup>200</sup>) were coupled on Actigel ALD beads (Sterogene). 15  $\mu$ l of beads were incubated with 5  $\mu$ g of either GST, GST-C-SH2, or GST-SH3 purified proteins (kind gift of Ivan Gout) for 4 h at 4 °C. Precipitates were then extensively washed, resuspended in sample buffer, denatured by 3 min boiling, then run on 10% SDS-PAGE, and silver-stained.

**D-3 Phosphoinositide Labeling, Extraction, and HPLC Separation**— $2 \times 10^8$  transfected cells were labeled with 1 mCi of [<sup>32</sup>P]orthophosphate (8500–9120 Ci/mmol, DuPont NEN) as described (28, 29). Following the labeling procedure, cells were washed and stimulated by the addition of CD80<sup>+</sup>-L cells. Cell contact was achieved by low speed centrifugation in a microcentrifuge for 5 s. Phospholipids were extracted as described (29, 30). The samples were deacylated and analyzed by anion exchange high performance liquid chromatography (HPLC) as described (28).

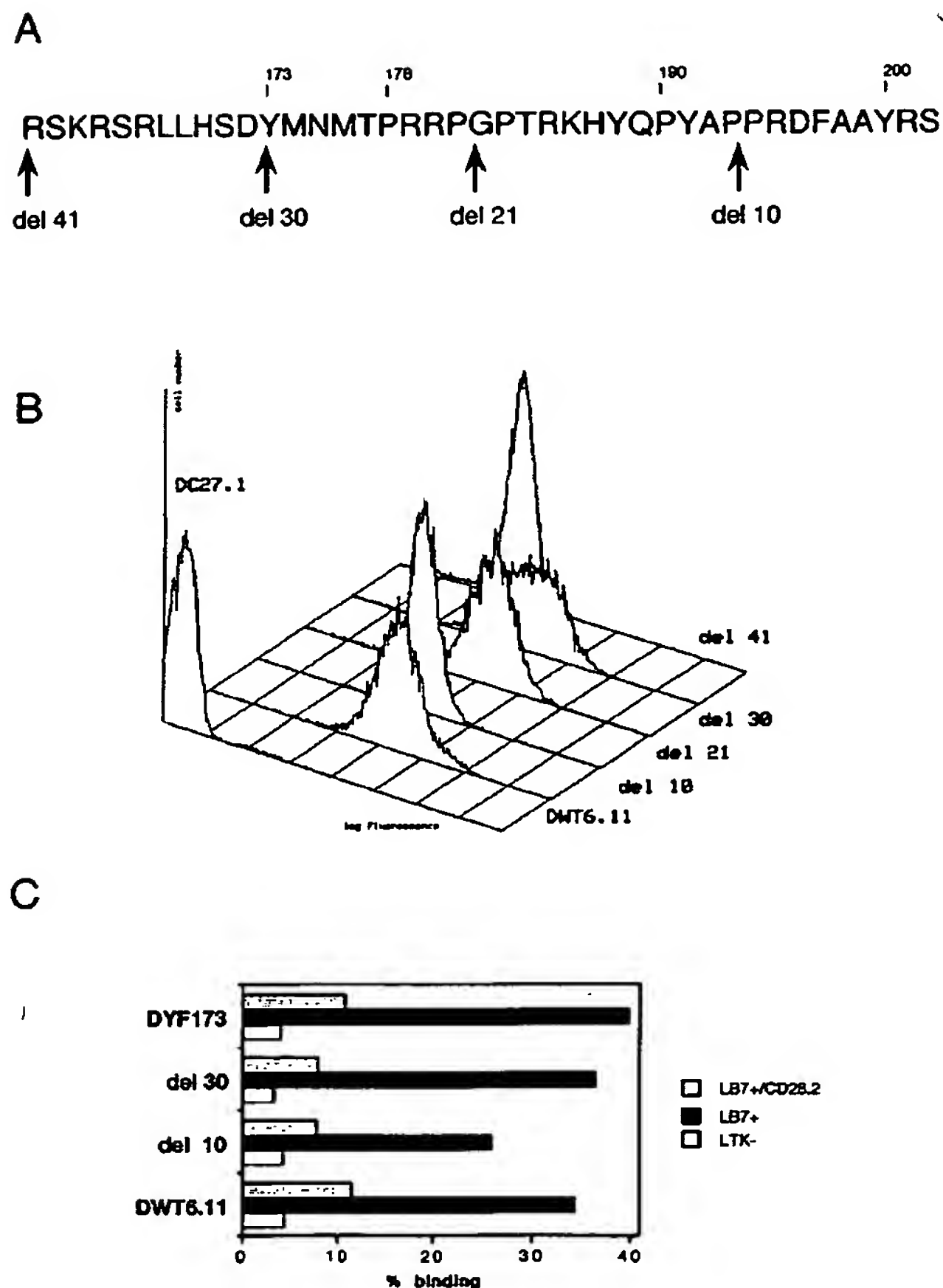
#### RESULTS

**Binding of B7/CD80-transfected Cells and CD28 mAbs to CD28 Intracytoplasmic Deletion Mutants Transfected into a Murine Hybridoma**—The high conservation of the CD28 cytoplasmic domain among various species (human, rat, mouse, chicken) suggested its major role in signal transduction. Nonetheless, in the absence of a recognizable catalytic domain within these sequences, an indirect coupling via adaptor molecules was suspected. We truncated 10 (*del 10*), 21 (*del 21*), 30 (*del 30*) C-terminal residues, respectively, or the whole intracellular domain (*del 41*) (Fig. 1A) and stably transfected these various constructs into the murine T-cell hybridoma DC27.1. Fig. 1B shows CD28 expression profiles for one clone representative of each transfection, after staining with the CD28.2 mAb. Deletion of 10, 21, 30, or 41 residues did not prevent surface expression of the transfected molecule (Fig. 1B), but for the *del 41* mutant, the mean fluorescence intensity was 7-fold lower than that observed for wild type CD28 (49 and 370, respectively). We also tested these cells for staining with a panel of 5 distinct mAbs: CD28.1, CD28.4, CD28.5, CD28.6, and 248 identifying at least 4 distinct epitopes on the CD28 molecule (25), as well as for binding of a B7-Ig fusion protein, and they all stained the various CD28 deletion mutants (not shown).

CD28 is an adhesion molecule since CD28/CD80 interaction allows cell adhesion (31). Using L cells transfected with human CD80, we show that wild type CD28-expressing cells bound to huCD80<sup>+</sup> cells (LB7<sup>+</sup>, 34.5% of binding) but not to untransfected cells (LTK<sup>−</sup>). In addition, this binding was inhibited by the addition of the human mAb CD28.2 (Fig. 1C). The *del 10* and *del 30* transfected mutants were still able to bind huCD80<sup>+</sup>-L cells with almost similar efficacy to wild type CD28. We previously reported the involvement of the tyrosine residue at position 173 in the activation of the PI 3-kinase pathway (12). Fig. 1C shows that this mutation did not affect CD28/CD80 interaction. Altogether, these data indicate that all deleted CD28 molecules still bind CD28 mAbs and B7-Ig and, in addition, are equally able to mediate the CD28/CD80 interaction showing that their extracellular structure was not modified.

*The 10 C-terminal Residues of CD28 Are Required for IL-2*

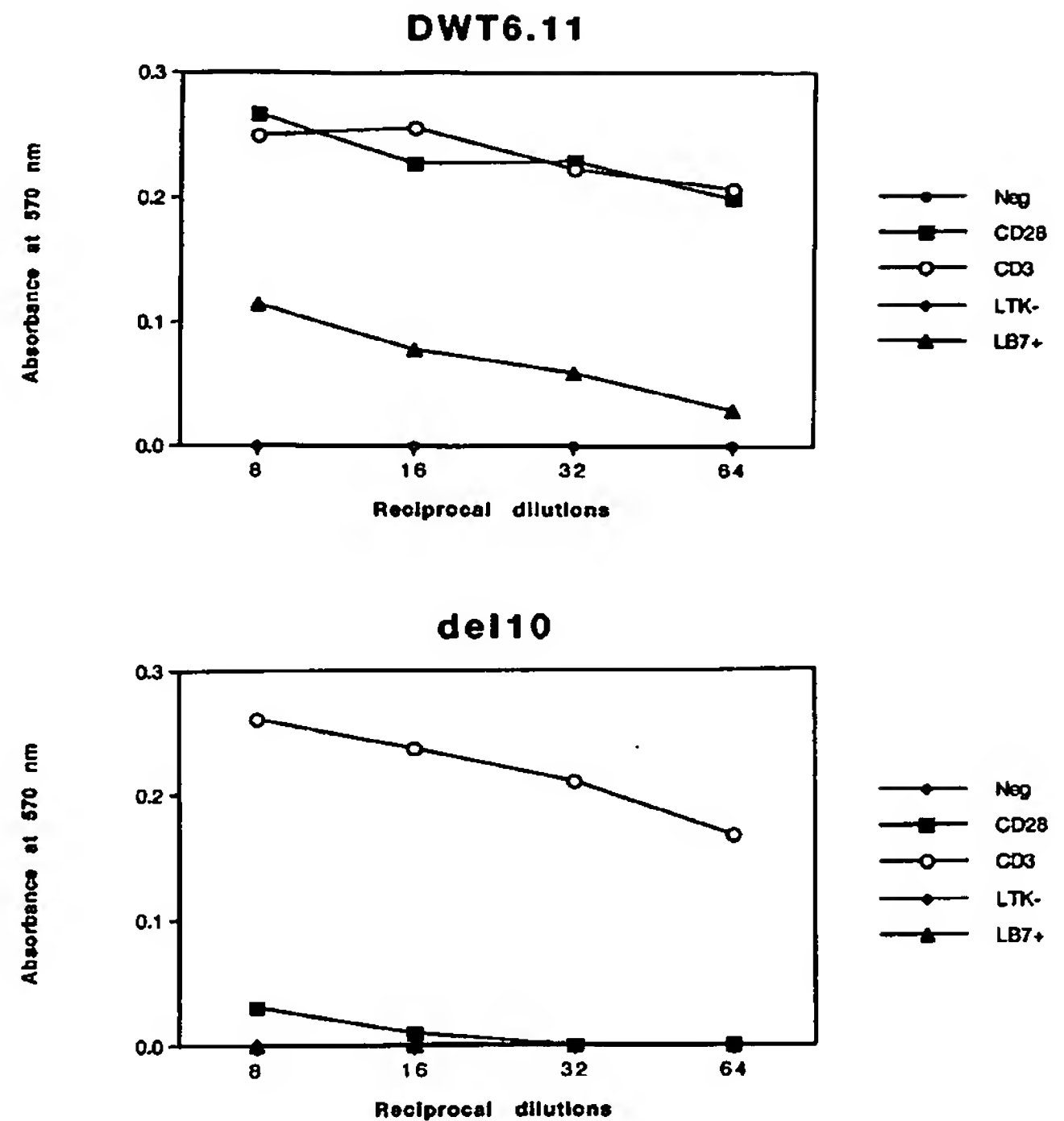
<sup>2</sup> A. Truneh, manuscript in preparation.



**FIG. 1. Intracytoplasmic truncations of the human CD28 molecule.** A, deletion mutants were produced by replacing original codons by stop codons (arrows) using oligonucleotide-directed mutagenesis. Sequencing of mutated molecules before transfection was performed according to the classical dideoxy method. B, one clone representative of each transfection (wild type or deleted CD28 molecules) was analyzed by flow cytometry after staining with the CD28.2 mAb. These fluorescence histograms were compared with staining of the untransfected murine T-cell hybridoma, DC27.1. C, adhesion assay was performed using untransfected (*LTK*<sup>-</sup>) or CD80<sup>+</sup>-L cells in the absence (*LB7*<sup>+</sup>), or presence of the CD28.2 mAb (*LB7*<sup>+</sup>/CD28.2).

**Secretion**—We have previously shown that a point mutation of the tyrosine residue at position 173 abolished both PI 3-kinase binding and IL-2 secretion in a murine T-cell hybridoma transfected with human CD28 (12). Here we investigated whether other regions of the CD28 cytoplasmic domain were required for late events of activation. We therefore tested whether CD28 stimulations (huCD80<sup>+</sup>-L cells or CD28 mAb in combination with CD3 mAb) could induce IL-2 secretion in transfected cells. Fig. 2 shows that both stimulations induced IL-2 secretion in cells transfected by the wild type CD28 construct (*upper panel*), while IL-2 production was severely altered in del 10 cells whatever CD28 stimuli was used (*lower panel*). Similar data were obtained with cells transfected by molecules truncated by 21, 30, and 41 residues (not shown). By contrast, cross-linked CD3 stimulation resulted in strong IL-2 secretion in all these clones.

**Coupling of Deleted CD28 Molecules to Phosphatidylinositol 3-Kinase and Tyrosine Kinases**—We and others have shown previously that, upon stimulation, the human CD28 molecule was able to associate with a PI 3-kinase activity. This association involves the SH2 domains of the p85 subunit interacting with a YNMN motif present in the CD28 cytoplasmic domain.



**FIG. 2. Function of wild type and deleted CD28 molecules.** Transfected cells were stimulated by cross-linked CD5 (closed circles) or CD3 (open circles) mAbs as negative and positive controls, respectively. CD28 stimulations were performed with CD80<sup>+</sup>-L cells (triangles), or soluble CD28 mAb in combination with soluble CD3 (closed squares). Supernatants were collected after 24 h of stimulation and titrated by serial dilutions for their ability to support proliferation of the IL-2-dependent cell line, CTLL-2. Results are expressed as  $A_{570\text{ nm}}$  obtained for each dilution of the supernatants and correspond to the proliferation of CTLL-2 as assessed by the cell growth determination 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide reagent. Stimulation by soluble CD28 or CD3 mAbs on their own did not induce a significant IL-2 production.

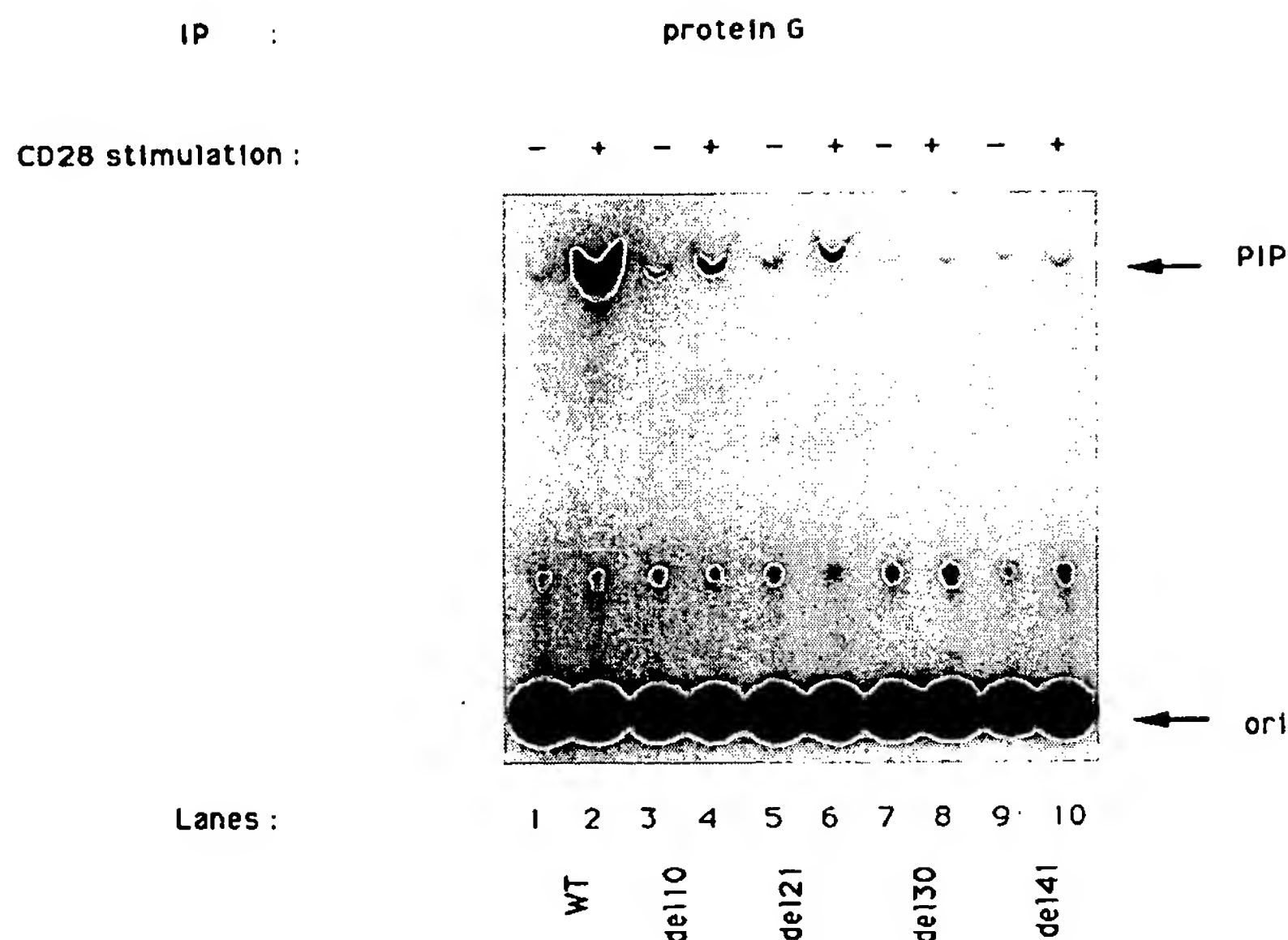
Here we tested gradually truncated CD28 molecules for their ability to associate with p85 and PI 3-kinase. Upon CD28 mAb stimulation PI 3-kinase activity associated with the wild type CD28 molecule (Fig. 3, *lane 2*), while a deletion of 30 C-terminal residues including the p85 binding site completely abolished this coupling (*lane 8*). Interestingly a deletion of only the 10 last residues was also able to inhibit 90% of the PI 3-kinase activity coupled to the CD28 molecule (*lane 4*).

Since the total immunoprecipitable PI 3-kinase activity was equivalent in all these cells (data not shown), the observed defect in PI 3-kinase activity could be explained either by the inability of truncated molecules to activate the enzyme or by their failure to associate with its p85 adaptor subunit. p85 Western blotting of CD28 immunoprecipitates revealed that deletion of 10 C-terminal residues decreased the CD28/p85 association by more than 90% while a deletion of 30 C-terminal amino acids including residues 173–176 completely abolished it (Ref. 12 and data not shown).

We also examined the ability of other transducing pathways to associate with CD28 deletion mutants. A rise in  $\text{Ca}^{2+}$  reflecting PLC $\gamma$ 1 activation was detected in cells expressing either wild type CD28 or del 10 mutant upon stimulation by CD3, as well as by CD28 mAbs (data not shown). CD28 and CD3 stimulations induce the tyrosine phosphorylation of specific substrates (17, 22, 23). A 2-min stimulation of both WT and del 10 transfected cells by CD3 mAbs induced the tyrosine phosphorylation of several substrates, the two most prominent



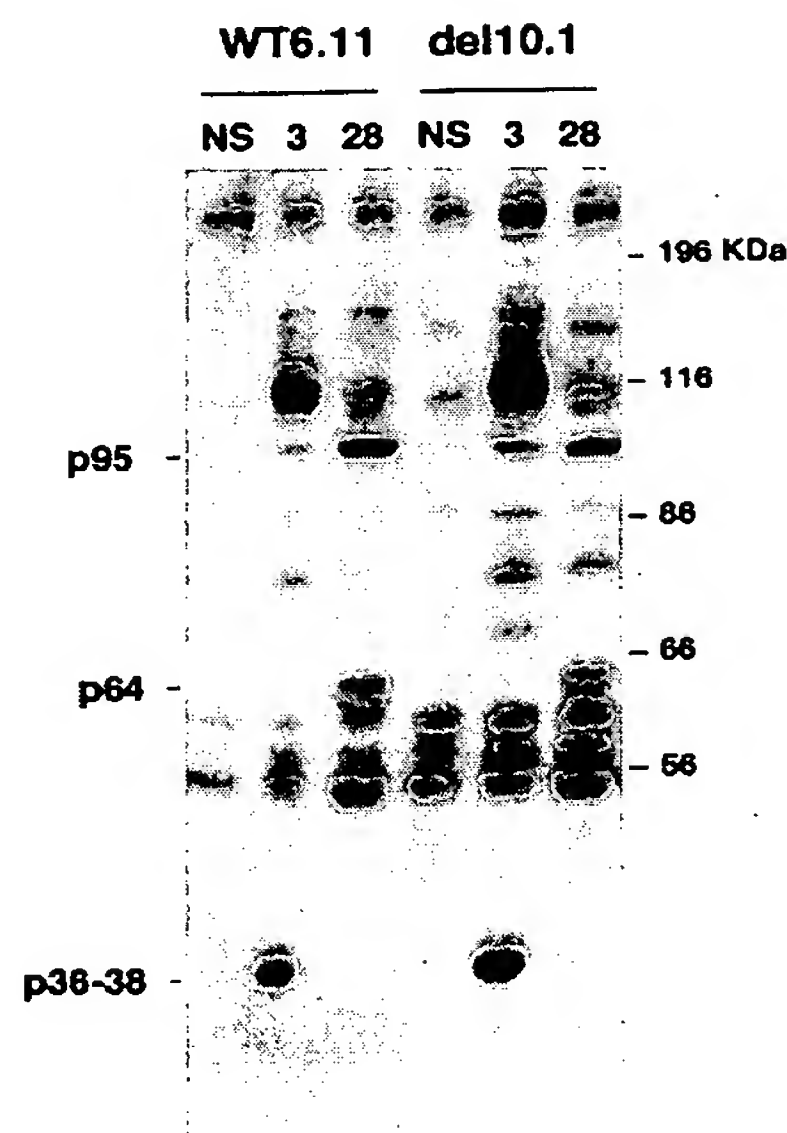
**FIG. 3. Association of CD28 mutants with phosphatidylinositol 3-kinase.** Each transfected clone, left unstimulated (lanes 1, 3, 5, 7, and 9) or stimulated with the CD28.2 mAb (lanes 2, 4, 6, 8, and 10), was tested for its ability to associate with PI 3-kinase. Immunoprecipitation of CD28 molecules was performed using protein G, and measure of PI 3-kinase activity associated with the various CD28 mutants was performed as described under "Experimental Procedures."



bands corresponding to molecular masses of 100 and 36 kDa (Fig. 4, lanes 2 and 5 (32, 33)). CD28 stimulation led to a strong phosphorylation of two proteins of 95 and 64 kDa, the former being *vav*.<sup>3</sup> As shown in Fig. 4, deletion of the 10 C-terminal amino acids did not prevent phosphorylation of these substrates.

**Binding of CD28 Peptides to p85 C-SH2 and SH3 Domains—**The cytoplasmic CD28 sequence contains two short proline-rich segments between residues 178–183 (PxxPxP) and 190–194 (PxxPP) which might serve as docking sites for the SH3 domain of p85 (34). Deletion of 10 C-terminal residues (amino acids 193–202) disrupts one of these proline-rich sequences, and this may account for the observed defect in p85 binding. To test this hypothesis and to determine if individual SH3 and C-SH2 domains of p85 could bind directly to the C-terminal part of CD28 *in vitro*, we tested whether a 17-mer peptide corresponding to residues 186–202 of CD28 could interact with recombinant SH2 and SH3 domain fusion proteins. As shown previously, interaction of p85 with the YNMN consensus site was strictly dependent upon tyrosine phosphorylation since a 15-mer phosphopeptide corresponding to residues 166–180 strongly bound the C-SH2 domain of p85 (Fig. 5, lane 4) while a nonphosphorylated form of the peptide did not (lane 2). In a non-phosphorylated form, peptide 186–202 did not interact with the C-SH2 domain (lane 6) while it did bind the p85 SH3 domain (lane 7). Interestingly, this p85 SH3 domain binding was decreased when peptide 186–202 was phosphorylated at position Tyr<sup>200</sup> (lane 10). Despite the absence of a consensus p85 binding site in the phosphopeptide 186–202, a weak binding of the p85 C-SH2 domain was observed however (lane 9).

**Involvement of Tyr<sup>200</sup> in CD28 Signaling—**To examine if the tyrosine residue at position 200 was involved in the PI 3-kinase pathway *in vivo*, we mutated it to phenylalanine and expressed the mutated construct in DC27.1 cells. Stable cell lines were analyzed for surface expression, CD80 binding, IL-2 secretion, p85 binding, and PI 3-kinase activation. Fig. 6A shows that point mutation of Tyr<sup>200</sup> → Phe<sup>200</sup> inhibited CD28-associated PI 3-kinase activity but did not abolish it (lane 4). Western blotting of p85 demonstrated that this impairment was due to



**FIG. 4. CD28-induced tyrosine phosphorylations in wild type and del 10 CD28 transfectants.** Wild type and del 10 transfectant cells were stimulated with CD3 (3, lanes 2 and 5) and CD28 (28, lanes 3 and 6) mAbs and goat anti-mouse Ig antiserum for 2 min, or left unstimulated (NS, lanes 1 and 4). Whole cell lysates were separated on SDS-PAGE, transferred onto polyvinylidene difluoride, and probed with anti-phosphotyrosine monoclonal antibody (4G10) as described under "Experimental Procedures."

a decrease in the quantity of CD28-associated p85 (not shown).

It has never been proven that PI 3-kinase association with CD28 was necessary for activation of the enzyme. We have therefore tested CD28-induced accumulation of D-3 phosphoinositides in the transfectants expressing wild type or mutated (Tyr<sup>173</sup>, Tyr<sup>200</sup>) CD28 molecules. In wild type transfectants, B7 ligation induces a transient accumulation of PtdIns(3,4,5)P<sub>3</sub> (Fig. 6B). However, point mutation of the Tyr<sup>173</sup> residue completely abolished CD28-induced PI 3-kinase activation as assessed by PtdIns(3,4,5)P<sub>3</sub> accumulation. In contrast, a point mutation of the Tyr<sup>200</sup> residue, however, only delayed and

<sup>3</sup> S. Klasen, F. Pagès, J. F. Peyron, D. Cantrell, and D. Olive, manuscript in preparation.



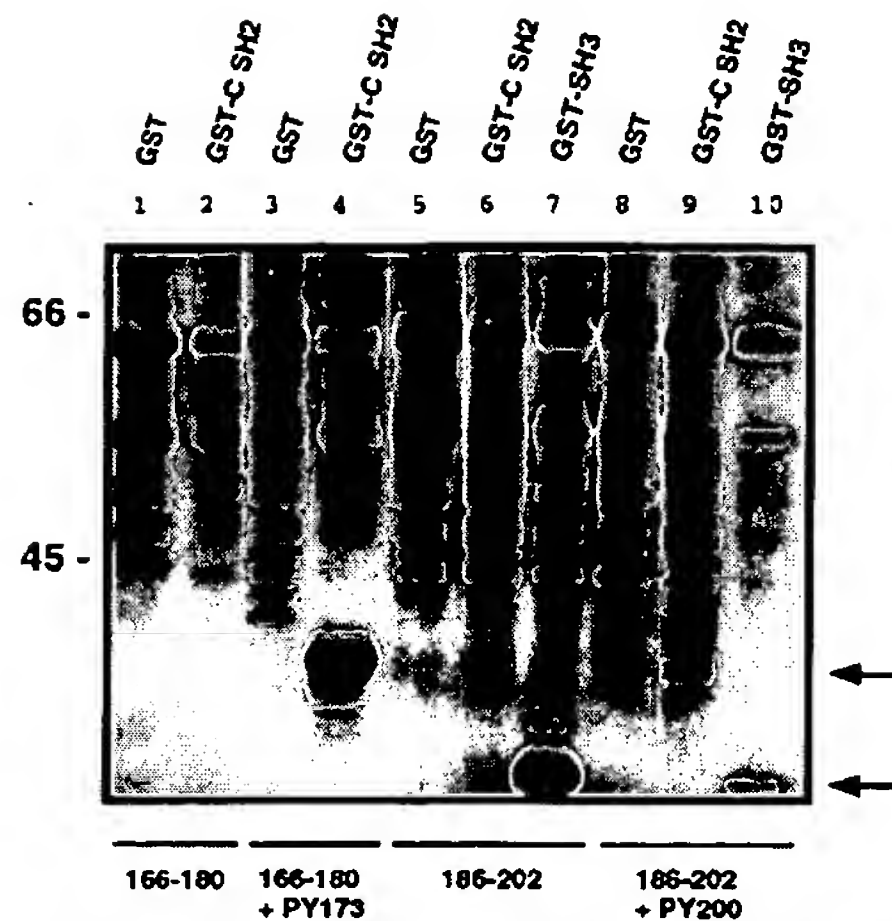


FIG. 5. Binding of CD28 peptides to purified p85 SH2 and SH3 domains. Peptides corresponding to CD28 residues 166–180 (lanes 1 and 2) including a phosphorylated Tyr<sup>173</sup> (lanes 3 and 4) and residues 186–202 (lanes 5–7) with phosphorylated Tyr<sup>200</sup> (lanes 8–10) were coupled on beads and used to precipitate recombinant GST (lanes 1, 3, 5, and 8), GST-C-SH2 (lanes 2, 4, 6, and 9), and GST-SH3 (lanes 7 and 10) p85 fusion proteins. Precipitates were run on 10% SDS-PAGE and revealed by silver staining.

attenuated its activation. Thus, the defect in the PI 3-kinase activation observed in the DYF200 transfectant occurs at the level of p85 association.

We also examined the ability of p85 C-SH2 domain fusion proteins to precipitate wild type or mutated CD28 molecules following ligation. Fig. 7A shows that upon ligation by CD80<sup>+</sup>-L cells, the GST-p85 C-SH2 fusion protein precipitated wild type CD28 (lane 2), while a point mutation of Tyr<sup>173</sup> strongly decreased this interaction (lane 8). Mutation of Tyr<sup>200</sup> did not prevent CD28 interaction with p85 C-SH2 domain (lane 5). Mutation of both Tyr<sup>173</sup> and Tyr<sup>200</sup> to phenylalanine abrogated most of the CD28 ability to be recognized by the C-SH2 domain of p85 (lane 11).

CD28 is tyrosine-phosphorylated upon activation (12, 21) and mainly on the Tyr<sup>173</sup> residue (21). Mutation of the Tyr<sup>200</sup> residue did not prevent CD28 phosphorylation, while point mutation of Tyr<sup>173</sup> strongly decreased it. The double mutant Tyr<sup>173-200</sup> lost most of its ability to be tyrosine-phosphorylated after stimulation (Fig. 7B). Hence, CD28 phosphorylation is further decreased by a double point mutation.

The tyrosine residue at position 200 is therefore involved in PI 3-kinase binding and activation. We have also tested its role in CD28 function. Fig. 8 shows that CD28 mAbs in combination with CD3 mAbs induced IL-2 secretion although to a lesser extent than wild type CD28. In contrast, stimulation by CD80<sup>+</sup>-L cells did not induce IL-2 secretion (Fig. 8) while binding to CD80<sup>+</sup>-L cells was retained (not shown).

#### DISCUSSION

In this report, we studied the structural requirements of the cytoplasmic domain of human CD28 for its signaling. For this analysis, the wild type CD28 molecule and various cytoplasmic mutants (deletion of 10, 21, 30, and 41 amino acids, or point mutation of Tyr<sup>200</sup> → Phe) were expressed into the murine T-cell hybridoma DC27.1. We have shown previously that transfection of the full-length CD28 cDNA in these cells allowed surface expression of functional molecules which induce either early or late events of T-cell activation (27). Flow cytometric analysis showed that deletion of 10, 21, or 30 residues did not affect the cell surface expression of CD28. Deletion of

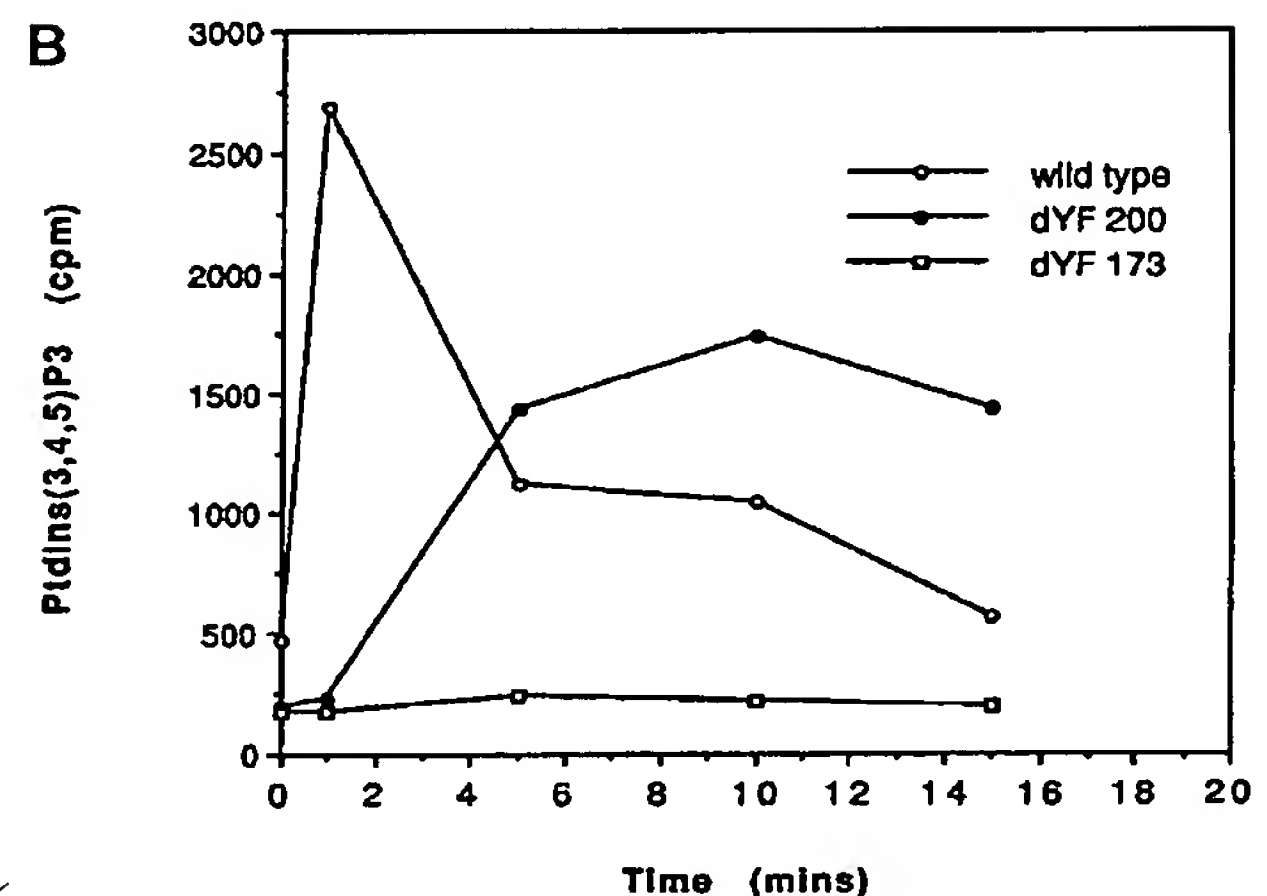
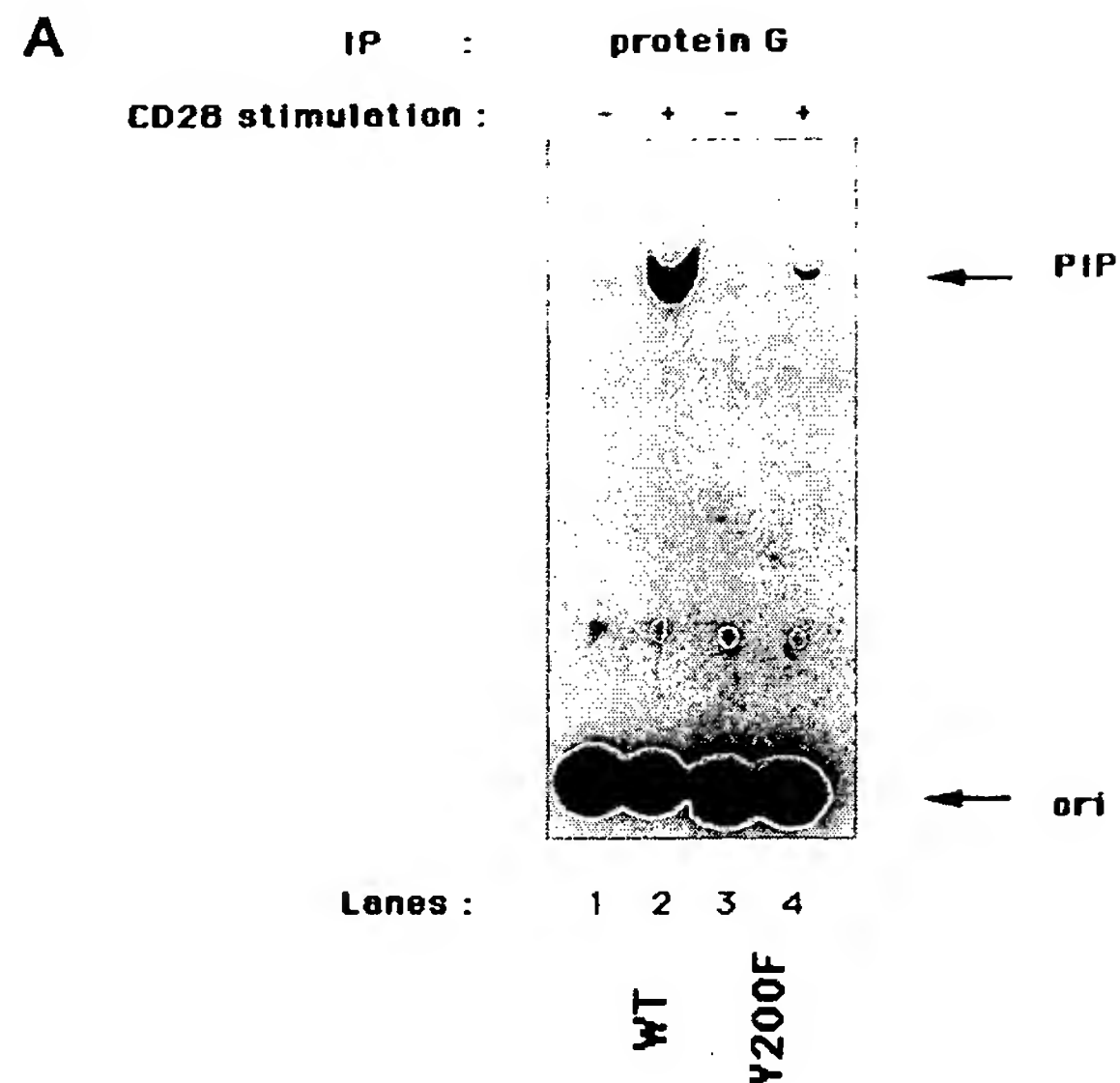


FIG. 6. Point mutation of the Tyr<sup>200</sup> → Phe<sup>200</sup>. A, wild type and mutated transfected cells were left unstimulated (lanes 1 and 3) or stimulated with the CD28.2 mAb (lanes 2 and 4) before immunoprecipitation with protein G. PI 3-kinase activity associated with CD28 was analyzed as in Fig. 3. B, PtdIns(3,4,5)P<sub>3</sub> generation. Wild type (open circles) or mutated (Tyr<sup>173</sup>, open squares; Tyr<sup>200</sup>, closed circles) CD28 transfectants were labeled with [<sup>32</sup>P]orthophosphate as described under "Experimental Procedures." Cells were then stimulated by the addition of CD80<sup>+</sup>-L cells. Phospholipids were extracted and analyzed by HPLC as described in Ref. 28.

the whole intracellular domain, however, impaired the expression of the construct. This observation has previously been reported for mutational analysis of the human CD2 molecule and could be explained by a partial instability of the molecule due to the removal of positively charged amino acids which are responsible for transmembrane stabilization (35).

After ligand binding and dimerization, many growth factor receptors phosphorylate several substrates on tyrosine residues leading to a cascade of signaling events. The antigen-binding T-cell receptor does not possess intrinsic enzymatic activity, and its coupling to the cellular signaling machinery is mediated by adaptor molecules. Mutagenesis studies of several molecules involved in T-cell functions (CD3  $\zeta$  chain, CD2) have identified cytoplasmic consensus motifs which couple these receptors to early events of T-cell activation. The ITAM motif

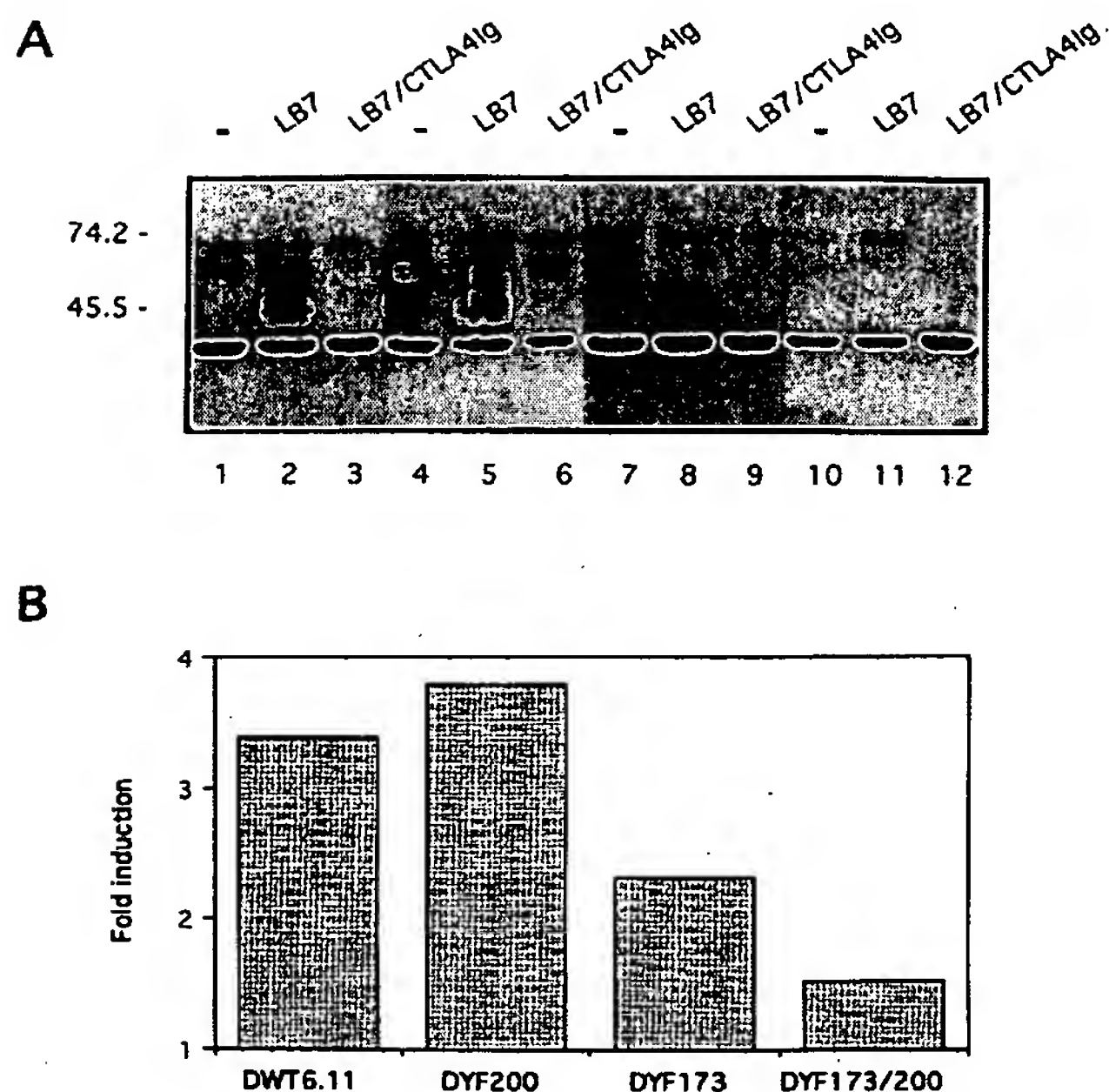


FIG. 7. Association of CD28 molecules with the C-SH2 domain of p85 and CD28 phosphorylation. *A*, cells transfected by wild type (lanes 1–3), Tyr<sup>200</sup> (lanes 4–6), Tyr<sup>173</sup> (lanes 7–9), or double-mutated Tyr<sup>173–200</sup> (lanes 10–12) CD28 molecules were stimulated by CD80<sup>+</sup>-L cells in the absence (lanes 2, 5, 8, and 11) or presence (lanes 3, 6, 9, and 12) of CTLA-4/Ig or left unstimulated (lanes 1, 4, 7, and 10). Whole cell lysates were precipitated using a p85 C-SH2 fusion protein. Precipitates were run on SDS-PAGE, and transferred onto polyvinylidene difluoride. Membranes were blotted using CD28.6 mAbs and revealed by ECL. *B*, blots were stripped and reprobed using an anti-phosphotyrosine mAb, and bands were quantified using the BioImage System (Millipore). Fold induction corresponds to the ratio of CD28 phosphorylation obtained in cells stimulated by CD80<sup>+</sup>-L cells versus unstimulated cells.

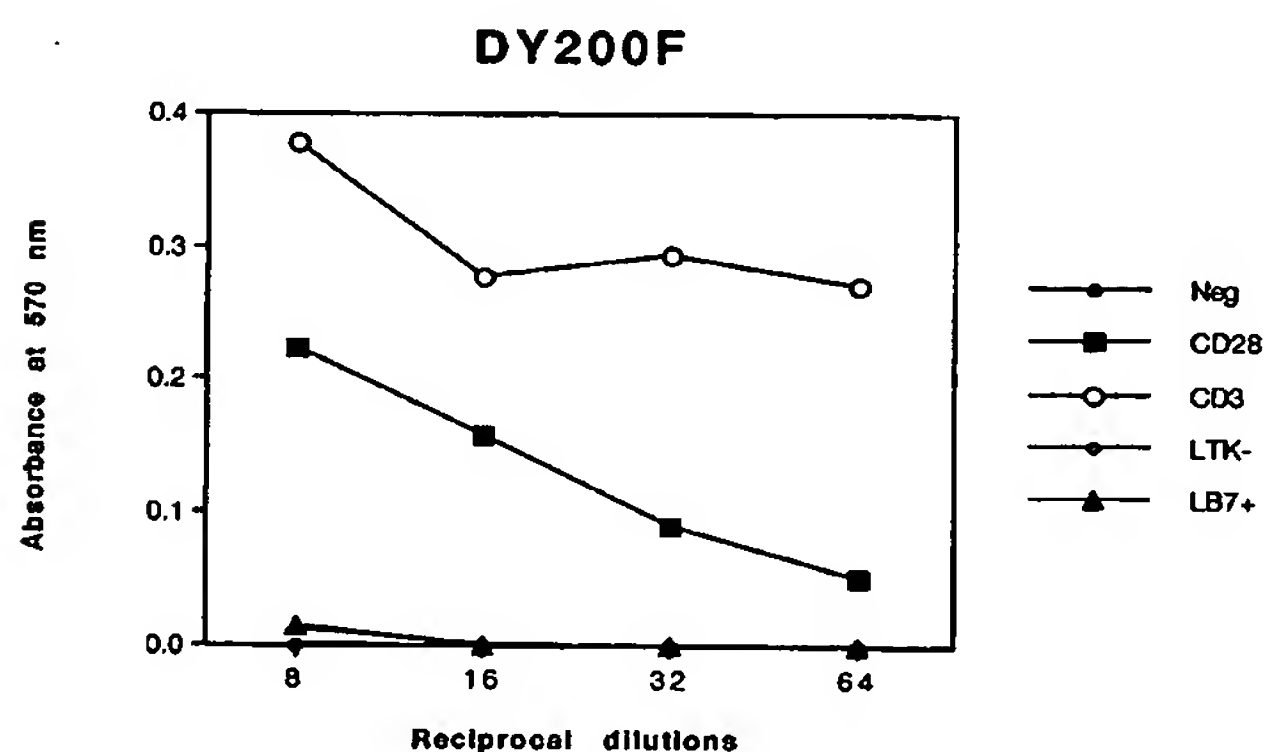


FIG. 8. IL-2 secretion in DYF200 transfectant. Cells transfected by mutated CD28 molecule were stimulated by cross-linked CD5 or CD3 mAbs (closed and open circles, respectively), by soluble CD28 mAbs in combination with soluble CD3 (closed squares) or by CD80<sup>+</sup>-L cells (triangles). Supernatants were collected after 24 h of stimulation and analyzed as described in Fig. 2.

(Yxx(I/L))<sub>2</sub> present in several subunits of the CD3 complex (36–38) couples the T-cell receptor to tyrosine kinase activation. Sequence comparison of CD28 with these molecules failed to identify any common motifs. Nonetheless, analysis of cytoplasmic sequences from human, mouse, rat, and chicken CD28 (7–10) showed high interspecies sequence similarity, suggesting a role for this domain in signal transduction. Functional

characterization of clones carrying mutations of CD28 confirms that the CD28 cytoplasmic domain plays a major role in signal transduction. We show that deletion of the 10 C-terminal amino acids severely impairs IL-2 secretion induced by a CD28 stimulation. This impairment was not merely due to a modification of CD28 extracellular structure, since all epitopes recognized by 6 different CD28 mAbs (25) were retained on the various deleted molecules, and since these transfectants were equally able to bind to B7-Ig and CD80-transfected L cells. Furthermore, cells carrying a deletion of 10 C-terminal residues were able to exhibit wild type levels of calcium mobilization as well as tyrosine phosphorylation of cellular substrates in response to CD28 stimulation. This suggests that the most C-terminal region of CD28 (residues 193–202) is crucial for the coupling of this receptor to IL-2 secretion. Interestingly, this region of CD28 is also involved in PI 3-kinase binding and activation. Together with the previously described loss of CD28 function following mutations of the PI 3-kinase binding site at residues Tyr<sup>173</sup> (12) and Met<sup>176</sup> (39), our data argue for the major role of this enzyme and/or its associated molecules in coupling the CD28 receptor to the cellular events leading to IL-2 secretion.

Upon ligand interaction, CD28 becomes tyrosine-phosphorylated and associates with p85 via a <sup>173</sup>YMN motif present in its cytoplasmic domain since a point mutation of Tyr<sup>173</sup> completely abolished p85 binding to CD28 (12, 13, 20). Recently, Raab *et al.* (21) have shown that p56<sup>lck</sup> and p59<sup>lyn</sup> can phosphorylate the Tyr<sup>173</sup> residue of CD28 *in vitro*. Interestingly, mutation of this residue did not completely abolish CD28 phosphorylation denoting the presence of other phosphorylation sites (21). Here we show that a deletion of 10 C-terminal residues greatly diminished the ability of CD28 to bind PI 3-kinase without affecting other signaling pathways such as PLCγ1 activation and tyrosine phosphorylations. Within this region, we have further identified two putative motifs involved in p85 binding. The first is a short proline-rich region (residues 190–194), and the second a tyrosine residue at position 200. We mutated this tyrosine residue (Tyr<sup>200</sup>) and confirmed its involvement in p85 binding and PI 3-kinase activation *in vivo*. Interestingly, deletion of the last 10 amino acids and point mutation of Tyr<sup>200</sup> only decreased PI 3-kinase binding to CD28. A low, but detectable, amount of the p85 still associated with mutated CD28. Furthermore, *in vitro* binding experiments showed that while the binding of the p85 SH2 domain to this Tyr<sup>200</sup> residue was dependent upon its phosphorylation, it was weak compared to binding to the <sup>173</sup>YMN motif. This observation was not surprising since Tyr<sup>200</sup> is not located within a consensus binding site for SH2 domains of p85 (YxxM, Ref. 40). Two alternative non-consensus binding sites, YVXV (41) and YVNA (42), have also been described as novel p85 recognition motifs in the tyrosine kinase receptors HGF-R and Flt-1, respectively.

The results we present here demonstrate that two regions in the intracytoplasmic domain of CD28 are involved in PI 3-kinase binding, one corresponding to the consensus p85 binding site <sup>173</sup>YMN and another one at the C terminus of the molecule (residues 193 to 202) including tyrosine 200 within a non-consensus p85 binding site. Although the CD28 <sup>173</sup>YMN motif is sufficient to associate with p85 since individual N- or C-SH2 fusion proteins can coprecipitate CD28 after CD28-B7 interaction and a 15-mer CD28 peptide including phosphorylated Tyr<sup>173</sup> precipitates PI 3-kinase from cell lysate (not shown), we propose that the two SH2 domains of p85 act in concert to associate with two distinct tyrosine residues of CD28 *in vivo*. The first is present within the consensus sequence <sup>173</sup>YMN and the other at position 200 is a non-consensus



binding site. This additional domain could either increase the affinity of p85/CD28 interaction or, alternatively, it could bind an adaptor molecule which interacts with p85. Although Tyr<sup>173</sup> is the major phosphorylation site in the CD28 cytoplasmic domain, our data support the hypothesis that Tyr<sup>200</sup> is also phosphorylated upon CD28 stimulation even though we do not directly demonstrate it. The observation that CD28 phosphorylation is further decreased by a double point mutation suggests that either the level of Tyr<sup>200</sup> phosphorylation is too weak to be detected in the presence of Tyr<sup>173</sup> phosphorylation, or, alternatively, that Tyr<sup>173</sup> plays a role in Tyr<sup>200</sup> phosphorylation, for instance by recruiting a tyrosine kinase. A third hypothesis is that Tyr<sup>200</sup>, although not being a direct target for phosphorylation, may be involved in the phosphorylation of other tyrosine residues of CD28 (residues 188 and 191).

Our *in vitro* binding experiments also showed that a nonphosphorylated peptide corresponding to the 17 C-terminal amino acids of CD28 also interacts with a purified GST-SH3 domain, probably through an interaction with the short CD28 proline-rich segment at residues 190–194 (PxxPP). Unexpectedly, phosphorylation of this peptide at position Tyr<sup>200</sup> decreased SH3 binding. The functional significance of this is at present unknown, and the *in vivo* relevance of this interaction has not been established since CD28 only associates with p85 after stimulation (12). This SH3/proline-rich interaction may, however, increase the affinity of p85 SH2 domains binding to CD28.

It is noteworthy that IL-2 secretion induced by a CD28 mAb costimulation was reduced, but not abolished, by mutation of Tyr<sup>200</sup> whereas it was completely impaired in the del 10 mutant. One explanation for the inability of deleted CD28 molecules to couple to the IL-2 secretory pathway is that a deletion of 10 C-terminal residues is sufficient to disrupt the CD28 intracytoplasmic structure. Nonetheless, this hypothesis is unlikely since other second messengers such as Ca<sup>2+</sup> rise and tyrosine phosphorylation of cellular proteins still occur upon CD28 activation. An alternative explanation is that, although PI 3-kinase is crucial for CD28 function, it is not the only transducing pathway involved in the coupling of CD28 to IL-2 secretion. Indeed, other enzymes such as sphingomyelinase have been reported to be involved in the CD28 costimulatory function (43), and their coupling to CD28 might also involve the C-terminal domain of the molecule. Consistent with this hypothesis is the observation that CD28 can function independently of PI 3-kinase, and that the CD28/PI 3-kinase association is not sufficient to mediate the full costimulatory function of CD28 (44).

Several reports have shown that SH3 domains of v-Src (45), p59<sup>lyn</sup> (46), or p56<sup>lck</sup> (47) could interact directly with the p85 subunit of PI 3-kinase, probably through proline-rich motifs identified at positions 88–97 and 299–309. This mechanism of PI 3-kinase coupling increases the complexity of possible interactions between transducing proteins. Yet another mechanism that might account for PI 3-kinase coupling to CD28 may involve indirect binding of p85 through an interaction with SH3 domains of one of these kinases previously coupled to CD28 via its C-terminal part. It has recently been shown that CD28 was phosphorylated by the Src-related protein-tyrosine kinases p56<sup>lck</sup> and p59<sup>lyn</sup> *in vitro*, and that this phosphorylation could increase the binding of p85, Grb-2, and ITK (21). In absence of a consensus binding site for SH2 domains of these kinases (Yxx(I/L), Ref. 40) in the intracytoplasmic domain of CD28, one of the questions remaining unanswered is how the Src kinases are recruited to the CD28 receptor after its stimulation.

**Acknowledgments**—We are indebted to Drs. Doreen Cantrell, Oreste Acuto, Patrick Auberger, Fergus McKenzie, Jan Domin, Ivan Gout, Jean Imbert, and Claude Mawas for their critical reading of the manu-

script, to G. Panayotou, P. Dubreuil, and R. Rottapel for their comments, and to Bernadette Barbarat for expert technical assistance.

## REFERENCES

- Allison, J. P. (1994) *Curr. Opin. Immunol.* 6, 414–419
- June, C. H., Bluestone, J. A., Nadler, L. M., and Thompson, C. B. (1994) *Immunol. Today* 15, 321–331
- Thompson, C. B. (1995) *Cell* 81, 979–982
- Chen, L., Ashe, S., Brady, W. A., Hellstrom, I., Hellstrom, K. E., Ledbetter, J. A., McGowan, P., and Linsley, P. S. (1992) *Cell* 71, 1093–1102
- Harding, F. A., McArthur, J. G., Gross, J. A., Raulet, R. H., and Allison, J. P. (1992) *Nature* 356, 607–609
- Townsend, S. E., and Allison, J. P. (1993) *Science* 259, 368–370
- Aruffo, A., and Seed, B. (1987) *Proc. Natl. Acad. Sci. U. S. A.* 84, 8573–8577
- Clark, G. J., and Dallman, M. J. (1992) *Immunogenetics* 35, 54–57
- Gross, J. A., St. John, T., and Allison, J. P. (1990) *J. Immunol.* 144, 3201–3210
- Young, J. R., Davison, T. F., Tregaskes, C. A., Rennie, M. C., and Vainio, O. (1994) *J. Immunol.* 152, 3848–3851
- August, A., and Dupont, B. (1994) *Int. Immunol.* 6, 769–774
- Pages, F., Ragueneau, M., Rottapel, R., Truneh, A., Nunes, J., Imbert, J., and Olive, D. (1994) *Nature* 369, 327–329
- Prasad, K. V. S., Cai, Y., Raab, M., Duckworth, B., Cantley, L., Shoelson, S. E., and Rudd, C. E. (1994) *Proc. Natl. Acad. Sci. U. S. A.* 91, 2834–2838
- Stein, P. H., Fraser, J. D., and Weiss, A. (1994) *Mol. Cell. Biol.* 14, 3392–3402
- Truitt, K. E., Hicks, C. M., and Imboden, J. B. (1994) *J. Exp. Med.* 179, 1071–1076
- Schneider, H., Cai, Y. C., Prasad, K. V., Shoelson, S. E., and Rudd, C. E. (1995) *Eur. J. Immunol.* 25, 1044–1050
- Nunes, J., Collette, Y., Truneh, A., Olive, D., and Cantrell, D. A. (1994) *J. Exp. Med.* 180, 1067–1076
- Su, B., Jacinto, E., Hibi, M., Kallunki, T., Karin, M., and Ben-Neriah, Y. (1994) *Cell* 77, 727–736
- August, A., Gibson, S., Kawakami, Y., Kawakami, T., Mills, G. B., and Dupont, B. (1994) *Proc. Natl. Acad. Sci. U. S. A.* 91, 9347–9351
- Hutchcroft, J. E., and Blierer, B. E. (1994) *Proc. Natl. Acad. Sci. U. S. A.* 91, 3260–3264
- Raab, M., Cai, Y. C., Bunnell, S. C., Heyeck, S. D., Berg, L. J., and Rudd, C. E. (1995) *Proc. Natl. Acad. Sci. U. S. A.* 92, 8891–8895
- Lu, Y., Granelli-Piperno, A., Bjornthal, J. M., Phillips, C. A., and Trevillyan, J. M. (1992) *J. Immunol.* 149, 24–29
- Vandenberghe, P., Freeman, G. J., Nadler, L. M., Fletcher, M. C., Kamoun, M., Turka, L. A., Ledbetter, J. A., Thompson, G. B., and June, C. H. (1992) *J. Exp. Med.* 175, 951–960
- Olive, D., Pages, F., Klasen, S., Battifora, M., Costello, R., Nunes, J., Truneh, A., Ragueneau, M., Martin, Y., Imbert, J., Birg, F., Mawas, C., Bagnasco, M., and Cerdan, C. (1995) *Fundamental Clin. Immunol.* 2, 185–197
- Nunes, J., Klasen, S., Franco, M. D., Lipcey, C., Mawas, C., Bagnasco, M., and Olive, D. (1993) *Biochem. J.* 293, 835–842
- Gunning, P., Leavitt, J., Muscat, G., Ng, S. Y., and Kedes, L. (1987) *Proc. Natl. Acad. Sci. U. S. A.* 84, 4831–4835
- Couez, D., Pages, F., Ragueneau, M., Nunes, J., Klasen, S., Mawas, C., Truneh, A., and Olive, D. (1994) *Mol. Immunol.* 31, 47–57
- Ward, S. G., Westwick, J., Hall, N. D., and Sansom, D. M. (1993) *Eur. J. Immunol.* 23, 2572–2577
- Jackson, T., Stephens, L., and Hawkins, P. T. (1992) *J. Biol. Chem.* 267, 16627–16636
- Stephens, L., Jackson, T., and Hawkins, P. T. (1993) *J. Biol. Chem.* 268, 17162–17172
- Linsley, P. S., Clark, E. A., and Ledbetter, J. A. (1990) *Proc. Natl. Acad. Sci. U. S. A.* 87, 5031–5035
- Sieh, M., Batzer, A., Schlessinger, J., and Weiss, A. (1994) *Mol. Cell. Biol.* 14, 4435–4442
- Buday, L., Egan, S. E., Viciana, P. R., Cantrell, D. A., and Downward, J. (1994) *J. Biol. Chem.* 269, 9019–9023
- Yu, H., Chen, J. K., Feng, S., Dalgarno, D. C., Brauer, A. W., and Schreiber, S. L. (1994) *Cell* 76, 933–945
- Chang, H., Moingeon, P., Lopez, P., Krasnow, H., Stebbins, C., and Reinherz, E. L. (1989) *J. Exp. Med.* 169, 2073–2083
- Letourneur, F., and Klausner, R. D. (1992) *Science* 255, 79–82
- Reth, M. (1989) *Nature* 338, 383–384
- Wegener, A. K., Letourneur, F., Hoeveler, A., Brocker, T., Luton, F., and Malissen, B. (1992) *Cell* 68, 83–95
- Cai, Y.-C., Cefal, D., Schneider, H., Raab, M., Nabavi, N., and Rudd, C. E. (1995) *Immunity* 3, 417–426
- Songyang, Z., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnoffsky, S., Lechleider, R. J., Neel, B. G., Birge, R. B., Fajardo, J. E., Chou, M. M., Hanafusa, H., Schaffhausen, B., and Cantley, L. C. (1993) *Cell* 72, 767–778
- Ponzetto, C., Bardelli, A., Maina, F., Longati, P., Panayotou, G., Dhand, R., Waterfield, M. D., and Comoglio, P. M. (1993) *Mol. Cell. Biol.* 13, 4600–4608
- Cunningham, S. A., Waxham, M. N., Arrate, P. M., and Brock, T. A. (1995) *J. Biol. Chem.* 270, 20254–20257
- Boucher, L.-M., Wiegmann, K., Futterer, A., Pfeiffer, K., Machleidt, T., Schutze, S., Mak, T. W., and Kronke, M. (1995) *J. Exp. Med.* 181, 2059–2068
- Truitt, K. E., Shi, J., Gibson, S., Segal, L. G., Mills, G. B., and Imboden, J. B. (1995) *J. Immunol.* 155, 4702–4710
- Liu, X., Marengere, L. E. M., Anne Koch, C., and Pawson, T. (1993) *Mol. Cell. Biol.* 13, 5225–5232
- Prasad, K. V. S., Janssen, O., Kapeller, R., Raab, M., Cantley, L. C., and Rudd, C. E. (1993) *Proc. Natl. Acad. Sci. U. S. A.* 90, 7366–7370
- Vogel, L. B., and Fugita, D. J. (1993) *Mol. Cell. Biol.* 13, 7408–7417



DISSECTION OF THE HUMAN CD2  
INTRACELLULAR DOMAIN  
Identification of a Segment Required for Signal  
Transduction and Interleukin 2 Production

BY HSIU-CHING CHANG,\*<sup>§</sup> PHILIPPE MOINGEON,\*<sup>§</sup>  
PETER LOPEZ,<sup>‡</sup> HELEN KRASNOW,\*  
CHRISTOPHER STEBBINS,\* AND ELLIS L. REINHERZ\*<sup>||</sup>

*From the \*Laboratory of Immunobiology and <sup>‡</sup>Flow Cytometry Facility, Dana-Farber Cancer  
Institute and Departments of <sup>§</sup>Pathology and <sup>||</sup>Medicine, Harvard Medical School,  
Boston, Massachusetts 02115*

The 50-kD CD2 (T11) molecule, originally defined as the sheep erythrocyte receptor, plays an important role in T lymphocyte activation as well as in facilitating adhesion between T lymphocytes and their cognate partners (1-5). Perturbation of the extracellular domain of CD2 by its ligand, LFA-3, or certain anti-CD2 mAbs provides signals that synergize to augment TCR-mediated stimulation (6, 7) and initiates a rapid turnover in polyphosphoinositide accompanied by an increase in cytosolic free calcium concentration (8, 9). In addition, this same combination of anti-CD2 antibodies (with or without LFA-3) induces lymphokine production from, and clonal expansion of, resting T lymphocytes (10).

Recently, the primary structures of human and murine CD2 were deduced from cDNA and genomic cloning (11-17). The predicted structure in both species is a type I integral membrane protein. In man, the mature protein consists of a hydrophilic 185 amino acid extracellular segment, a hydrophobic 25 amino acid transmembrane segment and a 117 amino acid cytoplasmic domain. The highest degree of homology between human and murine species is found in the cytoplasmic domain (59% at the amino acid level) and in both species the cytoplasmic domain is unique with respect to its multiple proline (21%) and basic residues and by secondary structural predictions, suggesting that it has an extended nonglobular conformation. Thus, unusual but conserved features among species including man, mouse, and rat suggest an important role for the CD2 segment in signal transduction (11, 12, 18). For this reason, we have performed a series of deletion mutation studies on the CD2 cytoplasmic domain in the present report.

#### Materials and Methods

*Construction of Truncated and Mutated Human CD2 Molecules and Transfection into a Murine T Cell Hybridoma.* To construct truncated CD2 molecules, the human CD2 cDNA, PB2 (11),

---

This work was supported by National Institutes of Health grant RO1 AI-21226. H. Chang was supported by a fellowship from the Cancer Research Institute. P. Moingeon is a recipient of a fellowship from the Philippe Foundation and Fondation pour la Recherche Medicale (France).

was digested with restriction enzymes Hph I, Ban II, Fok I, Stu I, and Ava I, respectively, and blunted with T4 DNA polymerase. The Hph I and Ban II digests were ligated to the linker 5'-CTAAGGATCCTTAG-3', while the Fok I, Stu I, and Ava I digests were ligated to the linker 5'-TAAGGATCCTTA-3' to regenerate the last amino acid, introduce a termination codon, and provide a Bam HI restriction recognition site. Subsequently, the DNAs were inserted into the Bam HI site of DOL vector (kindly provided by Dr. Thomas Roberts, Dana-Farber Cancer Institute). The substitution mutants were generated by oligonucleotide-directed in vitro mutagenesis as previously described (19). The full-length human cDNA PB2 insert was subcloned into BamHI site of M13 mp18. The synthetic oligonucleotides used for mutagenesis were 5'-CAGGCACCTAGTGATGAGCCCCCGCCTCCT-3' for CD2 M271-2, which changes the wild-type sequence CATCGT (His-Arg) into GATGAG (Asp-Glu), and 5'-CCGCCTCCTGGAGATGAGGTTTCAGCACCAG-3' for CD2 M278-9, which changes the wild-type sequence CACCGT (His-Arg) into GATGAG (Asp-Glu). The Bam HI fragment of the replicative form of M13 mutant DNAs was subcloned into expressing vector DOL. The plasmids containing full-length or modified CD2 cDNAs were isolated and sequenced around the modified region by double-stranded sequencing before transfection by  $\text{Ca}^{2+}$  precipitation into  $\psi$ -2, a helper-free retrovirus packaging cell line. Both transiently expressed and permanent viral stocks were used to infect the murine T cell hybridoma, 3DO54.8 (kindly provided by P. Marrack, National Jewish Hospital) (20) in the presence of 8  $\mu\text{g}/\text{ml}$  polybrene (Aldrich Chemical Co., Milwaukee, WI). Procedures for the growth of  $\psi$ -2 cells, transfection of cells, harvest of virus, and infection of cells were performed as described (21). Selection was initiated 48 h after infection using 0.4 mg/ml G418 (Geneticin; Gibco, Grand Island, NY), and wells containing single colonies were expanded. The G418 resistant clones were screened by indirect immunofluorescence as described in the legend to Fig. 2, and positive clones were further sorted on the Epics V cell sorter. Cells were maintained in RPMI 1640 (Gibco Laboratories) medium supplemented with 10% heat-inactivated FCS (Flow Laboratories, McLean, VA), 50  $\mu\text{M}$  2-ME, 1 mM sodium pyruvate, 2 mM L-glutamine, 1% penicillin-streptomycin (Whittaker, M. A. Bioproducts, Walkersville, MD) and 0.4 mg/ml G418.

**Flow Cytometric Analysis.** For cytometric analysis, ascites were used at a 1:200 dilution and  $10^6$  cell incubated for 30 min at 4°C. After washing in RPMI 1640 with 2% FCS, bound antibodies were detected using a 1:40 dilution of fluorescein-coupled goat anti-mouse IgG as a second antibody (Meloy, Springfield, VA). 10,000 cells were analyzed per sample on an Epics V cell sorter. Histograms represent the number of cells (ordinate) vs.  $\log_{10}$  fluorescence intensity (abscissa).

**Immunoprecipitation of CD2 Molecules.**  $10\text{--}20 \times 10^6$  cells were surface labeled with 1 mCi  $^{125}\text{I}$  (IMS 30; Amersham Corp., Arlington Heights, IL) for 15 min at room temperature using the lactoperoxidase method (11). Cell lysates were prepared in RIPA buffer containing 0.15 M NaCl, 1% Triton X-100, 0.1% sodium deoxycholate, 1 mM NaF, 1 mM PMSF, and protease inhibitors. Cell lysates were precleared once with a formalin-fixed *Staphylococcus aureus* suspension, then with Affigel protein A beads (Pharmacia Fine Chemicals, Piscataway, NJ) coupled to an irrelevant antibody (anti-CD8; 21Thy2D3) before overnight incubation at 4°C with the anti-CD2 (3T4-8B5) antibody coupled to beads. Immunoprecipitates were extensively washed with RIPA buffer and run on a 10% SDS-PAGE after treatment with 5% 2-ME. The gel was dried and the autoradiograph was exposed for 2 wk at  $-70^\circ\text{C}$  with intensifying screens.

**Measurement of Cytosolic  $\text{Ca}^{2+}$  by Indo-1 Fluorescence.** Cytosolic  $\text{Ca}^{2+}$  concentrations were determined according to Grynkiewicz et al. (22). Briefly,  $2 \times 10^6$  cells were loaded for 45 min at 37°C with 2  $\mu\text{g}/\text{ml}$  of the acetylmethyl ester of indo-1 (Molecular Probes, Junction City, OR) in 200  $\mu\text{l}$  of RPMI 1640 plus 2% FCS. Cells were diluted 10-fold before analysis on an Epics V cell sorter. Upon  $\text{Ca}^{2+}$  binding, indo-1 exhibits changes in fluorescein emission wavelengths from 480 to 410 nm (22). The ratio of 410/480 nm indo-1 fluorescence was recorded vs. real time and expressed in arbitrary units. One arbitrary unit represents  $\sim 200$  nM  $[\text{Ca}^{2+}]_i$ . For each determination, the baseline was assayed by recording indo-1-loaded cells for 1 min. Anti-T11<sub>2</sub> (1old24C1) and anti-T11<sub>3</sub> (1mono2A6) ascites were added at a 1:100 final dilution. The  $\text{Ca}^{2+}$  ionophore A23187 (Sigma Chemical Co., St. Louis, MO) was added at a 1  $\mu\text{g}/\text{ml}$  final concentration.

**Modulation of CD3 Molecules.** To modulate the murine CD3 molecule from the surface CD2 FL, cells were incubated overnight with the 145 2C11-purified antibody (20  $\mu$ g/ml) at 37°C. Cells were washed twice and an aliquot of the cell was used to evaluate the effect of modulation on murine CD3 and human CD2 expression using a standard indirect immunofluorescence assay (with the 145 2C11 and 3T4-8B5 mAbs, respectively). Modulated cells were loaded with indo-1 and cytosolic  $[Ca^{2+}]_i$  was followed upon various stimuli. Under such conditions, CD2 FL unmodulated cells or cells modulated with the anti-human CD3 mAb (Leu4) showed a  $[Ca^{2+}]_i$  comparable to Fig. 3 a (not shown) when triggered via CD2.

**Proliferation of CTLL-20 Cells.** For quantitation of IL-2 production,  $10^5$  cells/well were incubated in 96-well round-bottomed plates for 24 h in the presence of either ovalbumin (1 mg/ml final concentration) plus  $10^5$  A20-11 B lymphoma cells or anti-T11<sub>2</sub> + anti-T11<sub>3</sub> (ascites 1:100) or culture medium. Incubation with the nonstimulatory combination of anti-T11<sub>1</sub> + anti-T11<sub>2</sub> antibodies did not induce any detectable IL-2 production, while stimulation with anti-T11<sub>2</sub> + anti-T11<sub>3</sub> antibodies resulted in clear IL-2 secretion by CD2 FL cells. Because addition of PMA (Sigma Chemical Co.) was found to induce a substantial increase in lymphokine production, 5 ng/ml final concentration of PMA was added to all experimental samples including the media control. Subsequently, supernatants were harvested and titrated in triplicate for their ability to support the growth of 10,000 CTLL-20 cells. Cultures were pulsed after 24 h incubation with 1  $\mu$ Ci [<sup>3</sup>H]thymidine per well and harvested after an additional overnight incubation at 37°C over glass fiber filters on a Mash apparatus. Filters were dried and counted after addition of scintillation fluid on a  $\beta$  counter. Results are expressed as mean of triplicate determinations of cpm of [<sup>3</sup>H]thymidine incorporated. Standard deviations were generally <5-10% and results are representative of five independent experiments.

### Results and Discussion

To precisely characterize the functional and structural relationship of the cytoplasmic domain of the human CD2 molecule (Fig. 1 a), a series of deletion mutation of the CD2 cDNA were produced, as shown in Fig. 1 b, encoding 98, 77, 43, 18, and -3 amino acids out of the 117 predicted cytoplasmic CD2 amino acid residues. Full-length cDNAs as well as modified cDNAs were obtained as described in Materials and Methods and inserted into the retrovirus expression vector DOL (Fig. 1 c) under the control of the MLV LTR promoter (24) and defective viruses were generated (21). The murine T cell hybridoma 3DO54.8 cell line specific for ovalbumin in the context of the H-2 (I-A<sup>d</sup>) molecule (20) and lacking the human CD2 was then infected with these defective retroviruses. G418-resistant clones were selected in the presence of 0.4 mg/ml G418 and analyzed for surface CD2 expression using an indirect immunofluorescence assay with anti-T11<sub>1</sub> and anti-T11<sub>2</sub> antibodies on an Epics V cell sorter. Multiple clones corresponding to each type of truncation and expressing clearly detectable levels of surface CD2 were selected for further characterization. Clones used in subsequent functional studies were designated based on the nature of their CD2 cDNA retroviral insert: CD2 FL resulted from retroviral infection with the full-length CD2 cDNA, while CD2  $\Delta$ C98, CD2  $\Delta$ C77, CD2  $\Delta$ C43, CD2  $\Delta$ C18, and CD2  $\Delta$ C-3 resulted from infection with retroviruses containing the entire extracellular and transmembrane segment of CD2 but only 98 or fewer of the 117 cytoplasmic residues. A representative pattern of reactivity with the anti-CD2 mAb is shown in Fig. 2 a. All of these transfectants express comparable levels of CD2 ( $\sim$ 5-10,000 copies/cell) except CD2  $\Delta$ C18, which expresses on the order of 50% the copy number. The CD2  $\Delta$ C-3 transfectant expressed levels of CD2 comparable to CD2  $\Delta$ C18. However, the surface expression of CD2 on  $\Delta$ C-3 was unstable and did



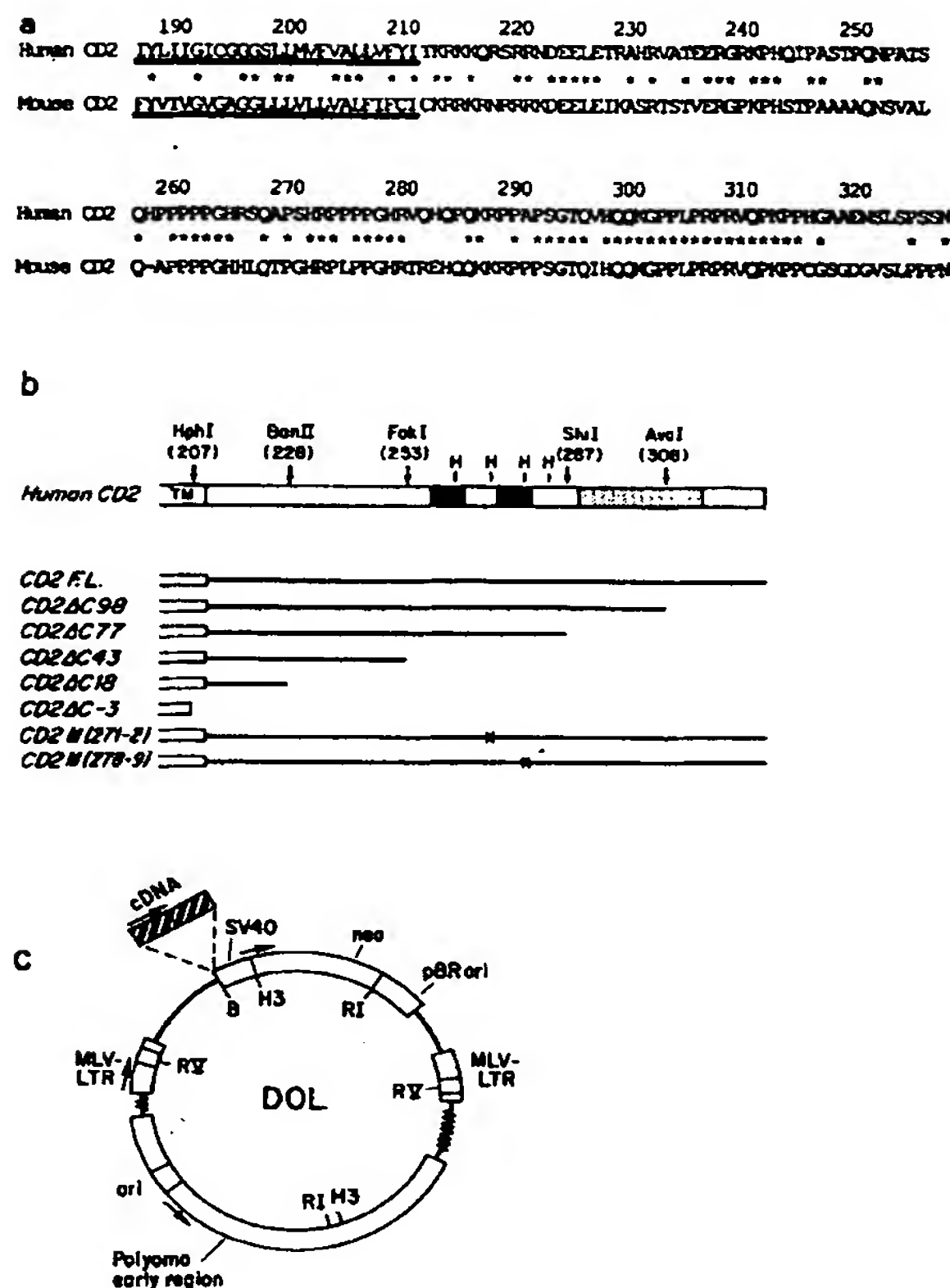


FIGURE 1. Production and analysis of the human CD2 cytoplasmic domain. (a) Comparison of predicted transmembrane and cytoplasmic domain sequences of human and mouse CD2. Amino acid residues are designated in single-letter code with the transmembrane regions underlined. Identical residues in human and mouse CD2 sequences are starred. (b) Schematic structure of the transmembrane and cytoplasmic regions of human CD2 and variant molecules. Constructs of full-length, deletion, and substitution mutants of CD2 are diagrammed. The region most conserved between human and mouse CD2 is stippled, and the two repeating PPPGHR segments are marked in black. The H denotes the histidine residues thought to form a putative binding site. The restriction sites that generate the truncated CD2 molecules are marked by arrows with numbers in parentheses corresponding to amino acid residues. (c) Structure of the DOL retroviral expression vector. Selected restriction sites on the vector shown are: B, Bam HI; RI, Eco RI; RV, Eco RV; H3, Hind III. The vector contains two promoters: MLV-LTR to drive the expression of the CD2 cDNA and the SV40 promoter to express the neomycin resistance gene (24).

not allow functional analysis (data not shown). For each clone, similar reactivities to those obtained with anti-T11<sub>1</sub> were found using the anti-T11<sub>2</sub> antibody, while none of the unactivated clones was stained by the anti-T11<sub>3</sub> antibody (data not shown).

To prove that individual clonal recipients of the truncated CD2 cDNAs express appropriately sized CD2 proteins, immunoprecipitation and SDS-PAGE analysis of the corresponding <sup>125</sup>I-labeled surface CD2 molecules was carried out. As expected, no human CD2 was immunoprecipitated from the murine 3DO54.8 cell line. In contrast, a 53-kD band was identified in SDS-PAGE analysis of clone CD2 FL. Note that the 70 kD band in anti-CD2 precipitations is unrelated to CD2 as it is present in the control immunoprecipitates. Furthermore, parallel analysis of CD2 protein expressed by CD2 ΔC98, ΔC77, ΔC43, and ΔC18 revealed that the molecular weights of surface CD2 (51, 47, 43, and 40 kD, respectively) correlated well with the expected truncations (Fig. 2 b). Thus, the truncated CD2 cDNAs in DOL direct protein synthesis of the variant CD2 forms on the surface of the murine 3DO54.8 cells.

Given that perturbation of the external domain of the CD2 molecule with a combi-

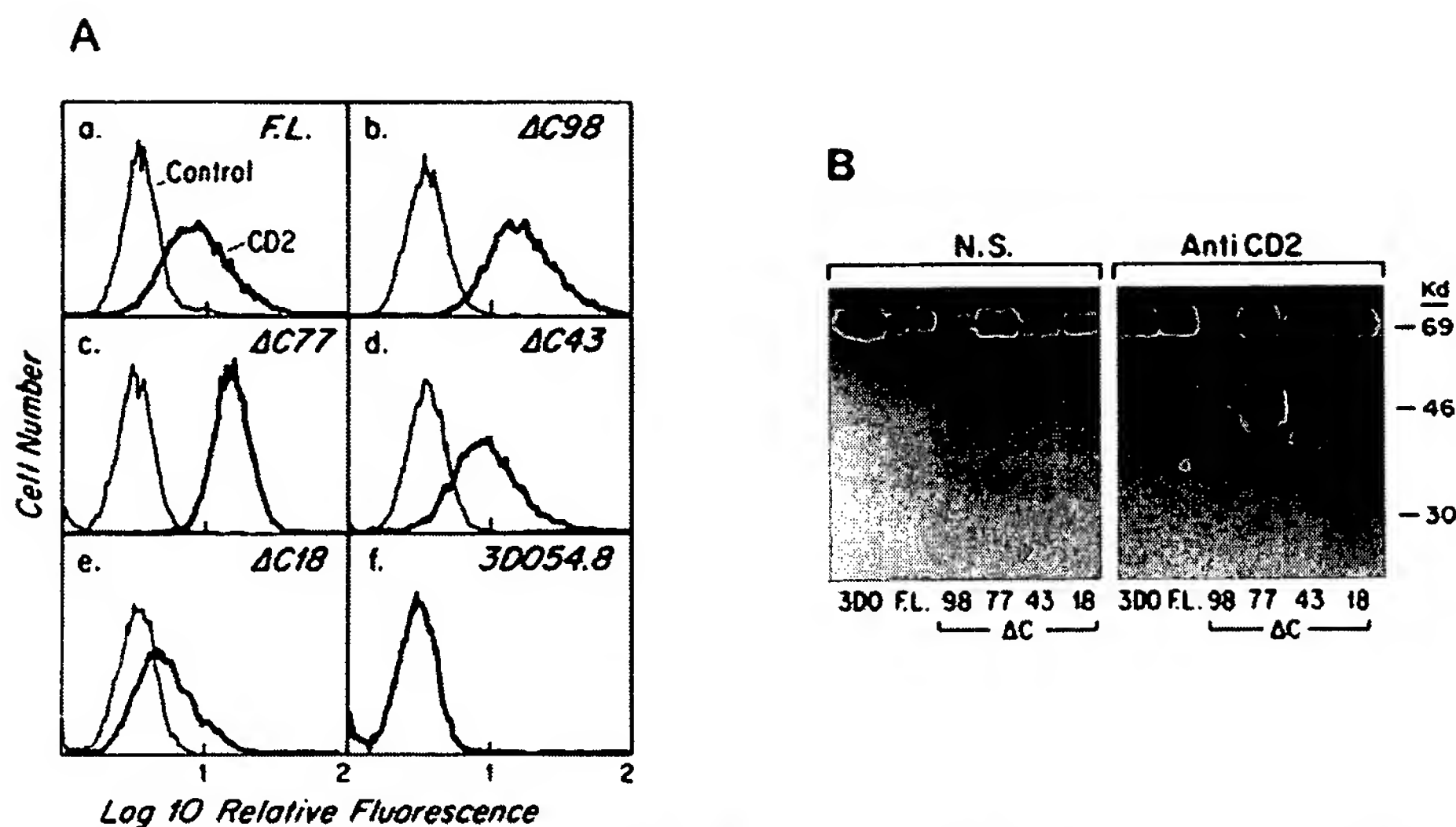


FIGURE 2. Expression of human CD2 molecules on murine T cells. (a) Flow cytometric analysis of CD2 expression on murine T cells. Indirect immunofluorescence assays were carried out using the anti-T11<sub>1</sub> mAb (3T4-8B5) (*thick line*) and compared with an irrelevant antibody (1HT4-4E5) (*thin line*) as background. (b) Immunoprecipitation of CD2 from lysates of iodinated cell lines. Immunoprecipitates were obtained from solubilized murine T cells using a nonspecific (NS) antibody (mouse anti-human CD8) or an anti-CD2 antibody directed against the T11<sub>1</sub> epitope and run under reducing conditions over an SDS 10% polyacrylamide gel (PAGE). The resulting autoradiogram is shown. A contaminant band of 70 kD (as well as material resolving at 200 kD, not shown) was regularly detected in the autoradiograms despite extensive preclearing, as previously reported by others (20).

nation of anti-T11<sub>2</sub> + anti-T11<sub>3</sub> results in a rapid rise in  $[Ca^{2+}]_i$  linked to IL-2 gene induction (1, 9), we examined whether such a mitogenic combination of mAbs was effective in stimulation of CD2 FL as well as CD2 ΔC cell lines. Fig. 3 a shows an analysis of alteration in  $[Ca^{2+}]_i$  after stimulation with anti-T11<sub>2</sub> + anti-T11<sub>3</sub> in various cell lines as measured with a calcium sensitive dye indo-1 (22) and flow cytometric analysis in real time. A clear rise in  $[Ca^{2+}]_i$  (~200 nM increment) was observed upon stimulation of CD2 FL, CD2 ΔC98 and CD2 ΔC77 cells. The calcium rise occurs within 2 min after adding the stimulating antibodies, most likely corresponding to the time required for expression of the T11<sub>3</sub> epitope after anti-T11<sub>2</sub> stimulation, a phenomena observed previously for human T lymphocytes (1). In contrast, CD2 ΔC43 and CD2 ΔC18 clones were not triggered by anti-T11<sub>2</sub> + anti-T11<sub>3</sub> antibodies. As expected, the nontransfected line 3DO54.8 was also not stimulated. Given that an immediate  $[Ca^{2+}]_i$  rise was observed after addition of the  $Ca^{2+}$  ionophore A23187 (1 μg/ml, final concentration), it is clear that cells were loaded with the fluorescent dye. These data establish that a significant rise in  $[Ca^{2+}]_i$  can be induced through human CD2 structures expressed on the membrane of murine T cells even in the absence of the COOH-terminal 40 amino acid residues of the CD2 cytoplasmic segment. The present results with human CD2 are consistent with two

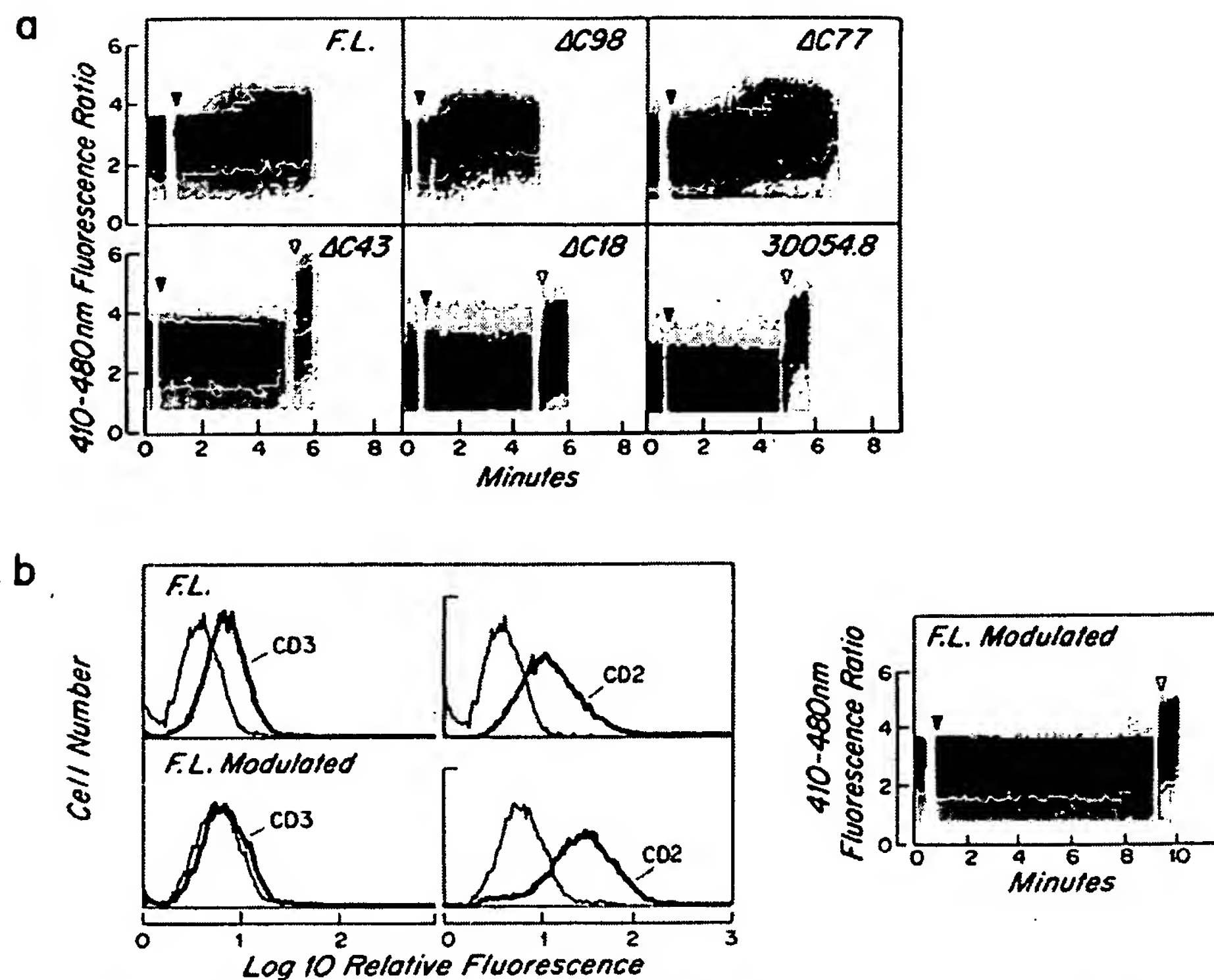


FIGURE 3. Elevation in intracellular free calcium mediated by CD2 stimulation. (a) Measurement of cytosolic  $\text{Ca}^{2+}$  by indo-1 fluorescence. Indo-1 fluorescence changes in real time histograms of T cell lines. Black and open arrows correspond to addition of anti-T11<sub>2</sub> + anti-T11<sub>3</sub> antibodies and calcium ionophore A23187, respectively. (b) Influence of CD3 modulation on induction of  $[\text{Ca}^{2+}]_i$  elevation triggered through the human CD2 molecule. Left panel represents indirect immunofluorescence detection of murine CD3 and human CD2 molecule, respectively, on the surface of unmodulated CD2 FL (top) and modulated CD2 FL (bottom). The latter was incubated with the 145 2C11 monoclonal antibody overnight (28). Right panel represents measurement of  $[\text{Ca}^{2+}]_i$  vs. time in modulated CD2 FL cells after stimulation with anti-T11<sub>2</sub> + anti-T11<sub>3</sub> antibodies (filled arrow) or calcium ionophore (open arrow).

independent reports. He et al. (25) demonstrated a requirement for the cytoplasmic tail of rat CD2 to trigger a rise in  $[\text{Ca}^{2+}]_i$  after transfection of rat CD2 in the human Jurkat T cell line. Bierer et al. (26) demonstrated a requirement for the cytoplasmic tail of human CD2 expressed in murine T cells for an IL-2 response to liposomes containing both HLA-DR and LFA-3 molecules.

Since human CD2 function in human T lymphocytes requires expression of the CD3-Ti  $\alpha/\beta$  complex (1, 3, 27), we examined whether the function of a human CD2 molecule within a murine cell line was linked to murine CD3-Ti. As shown in Fig. 3 b, modulation of the murine CD3 molecule by the hamster anti-mouse CD3 $\epsilon$  antibody (145 2C11) resulted in nearly complete loss of the CD3 molecule from the cell surface of CD2 FL after incubation for 16 h at 37°C, while CD2 expression



was unaltered or slightly increased. Such anti-CD3 modulated CD2 FL cells were no longer stimulated by anti-T11<sub>2</sub> + anti-T11<sub>3</sub> antibodies to increase  $[Ca^{2+}]_i$ . By contrast, the incubation of CD2 FL cells under similar conditions with anti-human CD3 antibody (Leu4) did not have any effect (Materials and Methods). This result suggests that regulation of the CD2 pathway by CD3-Ti is intact in the cellular model herein, further supporting the validity of this functional approach in the study of the CD2 intracellular domain. We cannot, of course, exclude the possibility that CD3 modulation leads to a generalized disruption of subsequent T cell activation.

We next examined if nuclear activation events including IL-2 induction and subsequent IL-2 secretion could be triggered through the human CD2 molecule in murine CD2 FL cells. To this end, clones were stimulated and IL-2 secretion into supernatants assayed using the IL-2-dependent CTLL-20 cells (29). As shown in Fig. 4a, supernatants from CD2 FL or 3DO54.8 cells stimulated with ovalbumin in the presence of the H-2(I-A<sup>d</sup>) expressing A20-11 B lymphoma are able to induce proliferation of the IL-2-dependent CTLL-20 cells in a comparable way. By contrast, the

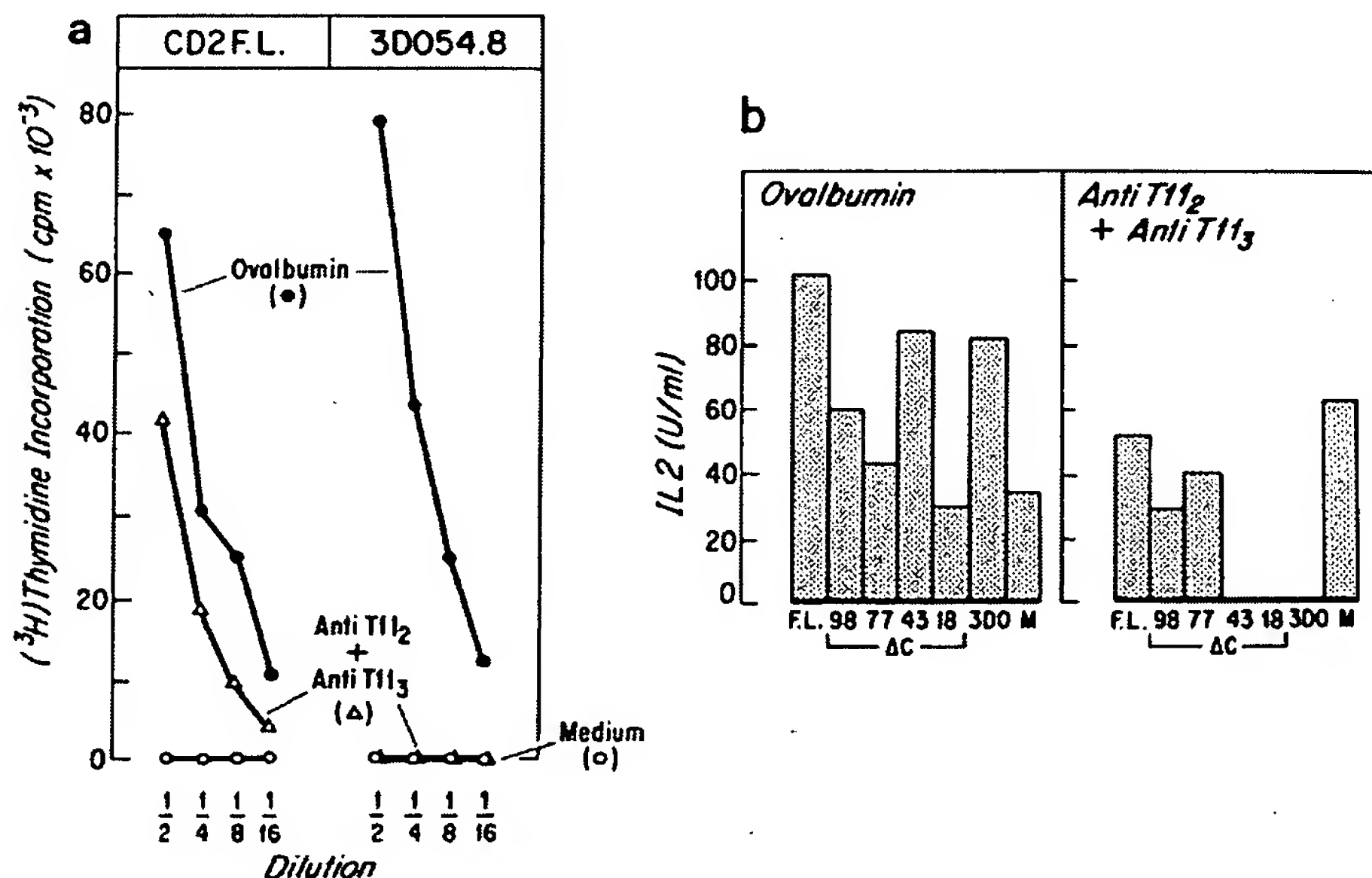


FIGURE 4. Stimulation of IL-2 production through the human CD2 molecule and its variants forms on murine T cells. (a) Effect of culture supernatants from CD2 FL or 3DO54.8 on proliferation of CTLL-20 cells. Serial twofold dilutions of culture supernatants from stimulated or unactivated murine cells were tested for their capacity to support the growth of the IL-2-dependent murine CTLL-20 cell line (29), as evaluated by  $[^3H]$ thymidine incorporation. (b) IL-2 production upon antigen (ovalbumin) or anti-T11<sub>2</sub> + anti-T11<sub>3</sub> stimulation. Under these experimental conditions, the limit of detection for IL-2 was  $<4$  U/ml. M represents the murine T cell line CD2 M271-2. Analysis of IL-2 production by two independent clones of the CD2 M278-9 provided evidence that such cells could produce significant amounts of IL-2 upon CD2 triggering (~50% of the amount secreted upon antigen stimulation). Experiments were performed as above, but precise quantification of IL-2 production was obtained by running culture supernatants in parallel to a titration curve of rIL-2 (Biogen Laboratories, Cambridge, MA).

combination of anti-T11<sub>2</sub> + anti-T11<sub>3</sub> antibodies is effective in inducing IL-2 production by CD2 FL cells but not the parental line 3DO54.8. Fig. 4 *b* shows that all of the cell lines tested including CD2 FL, 3DO54.8, and the CD2  $\Delta$ C series produce a high amount of IL-2 (ranging from 30 to 100 U/ml, corresponding to a clonal variation repeatedly observed) when stimulated with ovalbumin in the I-A<sup>d</sup> context. This result demonstrates the integrity of the IL-2 synthetic pathway in each clone. Perhaps more importantly, after stimulation with anti-CD2 antibodies, the clones CD2 FL, CD2  $\Delta$ C98, and CD2  $\Delta$ C77 produce comparable levels of IL-2, while transfectants CD2  $\Delta$ C43 and CD2  $\Delta$ C18, like the untransfected 3DO54.8, are not triggered through CD2. Analysis of five other independent clones expressing the  $\Delta$ C43 CD2 molecule clearly established that these cells are not triggered through human CD2 either to secrete detectable levels of IL-2 or to elevate  $[Ca^{2+}]_i$  (data not shown). Taken together, these data show that a full-length human CD2 molecule, as well as a CD2 molecule lacking 19 or 40 COOH-terminal amino acids from the cytoplasmic domain, is able to activate T lymphocytes after appropriate perturbation of the CD2 extracellular segment. Interestingly, the CD2  $\Delta$ C77 clones express human CD2 molecules lacking residues 289 to 316. The latter corresponds to the segment most conserved among human and murine molecules, with 24 out of 27 residues being identical (Fig. 1, *a* and *b*). Presumably, these conserved residues function in another facet of CD2 biology unrelated to IL-2 induction and/or secretion. In contrast, the  $\Delta$ C43 truncated molecules as well as shorter truncations are nonfunctional with respect to stimulating a rise in  $[Ca^{2+}]_i$  and IL-2 production.

These data provide unequivocal evidence that the CD2 cytoplasmic domain is involved in signal transduction and that one essential sequence of the cytoplasmic domain necessary for CD2-mediated activation is located between amino acids 253 to 287. This region contains four histidines at amino acid positions 264, 271, 278, and 282 and includes two tandemly repeated segments (PPPGHR, amino acids 260–265 and 274–279) (see Fig. 1, *a* and *b*). These histidine residues could represent a binding site for an ion, cyclic nucleotide, or other small regulatory molecule.

To evaluate the role of the histidine residues, substitution mutants of CD2 were produced by site-directed mutagenesis as described in Materials and Methods. Two different categories of mutants with a stable CD2 surface expression were obtained: CD2 M271-2 and CD2 M278-9 in which the positively charged histidine and arginine residues at position 271 and 272 or 278 and 279 were replaced by negatively charged aspartic acid and glutamic acid, respectively (Fig. 1 *b*). Functional characterization of these two mutants shows that IL-2 production can be induced by anti-T11<sub>2</sub> and anti-T11<sub>3</sub> antibodies to a level comparable to that of the CD2 FL clones (Fig. 4 *b* and data not shown). Thus, a putative "cage," requiring four histidine residues, is not necessary for this CD2 activation. Furthermore, since the mutations at position 278–279 alter the structure of the second repeat without affecting T cell activation, the more COOH-terminal repeat is not required. Therefore, it may be speculated that perhaps only one PPPGHR sequence is necessary for induction of IL-2 activity. This notion can be tested by further site-directed mutagenesis studies. Alternatively, the repeats and/or other histidine residues could function to regulate CD2 interaction with cytoskeletal, plasma membrane, cytosolic, nuclear, or other cellular components not examined by the present assays. It also remains to be determined if the CD2 cytoplasmic region influences the CD2 extracellular segment and

hence, interaction with LFA-3. The present model system will be useful at addressing these latter possibilities.

### Summary

To evaluate those residues in the 117 amino acids of the CD2 cytoplasmic domain required for transduction of T lymphocyte activation signals, a full-length human CD2 cDNA and a series of deletion and substitution mutants were inserted into the ovalbumin-specific, I-A<sup>d</sup>-restricted murine T cell hybridoma 3DO54.8 using a retroviral system. The resulting cells express surface CD2 protein and unlike the parental murine line, are reactive with murine anti-human CD2 antibodies. Anti-T11<sub>2</sub> plus anti-T11<sub>3</sub> antibody stimulation of cells expressing a full-length CD2 cDNA results in a characteristic rise in cytosolic-free calcium ( $[Ca^{2+}]_i$ ), and subsequent IL-2 secretion that accompany CD2 stimulation in human T lymphocytes. Transfectants expressing CD2  $\Delta$ C98 and CD2  $\Delta$ C77, partially deleted CD2 molecules containing the entire extracellular and transmembrane CD2 segments but only 98 and 77 amino acids of the cytoplasmic domain, respectively, are also activated by anti-CD2 mAbs. In contrast, clones expressing more severely truncated CD2 structures, CD2  $\Delta$ C43 and CD2  $\Delta$ C18, are not stimulated. These data show that the cytoplasmic domain plays an essential role in transduction of activation signals via CD2, and that the segment between amino acid residues 253 and 278 is necessary for activation. This region contains two tandem repeats of the sequence PPPGHR, thought to form part of a putative cationic site. Disruption of the latter by site-directed mutagenesis does not affect IL-2 gene induction, suggesting that only one of the repeats is required for activating this function of the CD2 molecule.

*Received for publication 30 November 1988 and in revised form 16 February 1989.*

### References

1. Meuer, S. C., R. E. Hussey, M. Fabbi, D. Fox, O. Acuto, K. A. Fitzgerald, J. C. Hodgdon, J. P. Protentis, S. F. Schlossman, and E. L. Reinherz. 1984. An alternative pathway of T cell activation: a functional role for the 50KD T11 sheep erythrocyte receptor protein. *Cell*. 36:897.
2. Siliciano, R. F., J. C. Pratt, R. E. Schmidt, J. Ritz, and E. L. Reinherz. 1985. Activation of cytolytic T lymphocytes and natural killer cell function through the T11 sheep erythrocyte binding protein. *Nature (Lond.)*. 317:428.
3. Brottier, P., L. Boumsell, C. Gelin, and A. Bernard. 1985. T cell activation via CD2 (T, gp50) molecules: accessory cells are required to trigger T cell activation via CD2-D66 plus CD2-9.6/T11 (1) epitopes. *J. Immunol.* 135:1624.
4. Selvaraj, P., M. L. Plunkett, M. Dustin, M. E. Sanders, S. Shaw, and T. A. Springer. 1987. The T lymphocyte glycoprotein CD2 binds the cell surface ligand LFA-3. *Nature (Lond.)*. 326:400.
5. Dustin, M. L., M. E. Sanders, S. Shaw, and T. A. Springer. 1987. Purified lymphocyte function-associated antigen 3 binds to CD2 and mediates T lymphocyte adhesion. *J. Exp. Med.* 165:677.
6. Yang, S. Y., S. Chouaib, and B. DuPont. 1986. A common pathway for T lymphocyte activation involving both the CD3-Ti complex and the CD2 sheep erythrocyte receptor determinants. *J. Immunol.* 137:1097.
7. Bierer, B. E., J. Barbosa, S. Herrmann, and S. J. Burakoff. 1988. The interaction of



- CD2 with its ligand, LFA-3, in human T cell proliferation. *J. Immunol.* 140:3358.
8. Alcover, A., M. J. Weiss, J. F. Daley, and E. L. Reinherz. 1986. The T11 surface glycoprotein is functionally linked to a calcium channel in precursor and mature T lineage cells. *Proc. Natl. Acad. Sci. USA.* 83:2614.
  9. Pantaleo, G., D. Olive, A. Poggi, W. J. Kozumbo, L. Moretta, and A. Moretta. 1987. Transmembrane signalling via the T11-dependent pathway of human T cell activation. Evidence for the involvement of 1,2-diacylglycerol and inositol phosphates. *Eur. J. Immunol.* 17:55.
  10. Hunig, T., G. Tiefenthaler, K. H. Meyer Zum Buschenfelde, and S. C. Meuer. 1987. Alternative pathway activation of T cells by binding of CD2 to its cell surface ligand. *Nature (Lond.)* 326:298.
  11. Sayre, P. H., H. C. Chang, R. E. Hussey, N. R. Brown, N. E. Richardson, G. Spagnoli, L. K. Clayton, and E. L. Reinherz. 1987. Molecular cloning and expression of T11 cDNAs reveal a novel receptor-like structure on human T lymphocytes. *Proc. Natl. Acad. Sci. USA.* 84:2941.
  12. Clayton, L. K., P. H. Sayre, J. Novotny, and E. L. Reinherz. 1987. Murine and human T11 (CD2) cDNA sequences suggest a common signal transduction mechanism. *Eur. J. Immunol.* 17:1367.
  13. Sewell, W. A., M. H. Brown, J. Dunne, M. J. Owen, and M. J. Crumpton. 1986. Molecular cloning of the T lymphocyte surface CD2 (T11) antigen. *Proc. Natl. Acad. Sci. USA.* 83:8718.
  14. Seed, B., and A. Aruffo. 1987. Molecular cloning of the CD2 antigen, the T cell erythrocyte receptor, by a rapid immunoselection procedure. *Proc. Natl. Acad. Sci. USA.* 84:3365.
  15. Diamond, D. J., L. K. Clayton, P. H. Sayre, and E. L. Reinherz. 1988. Exon-intron organization and sequence comparison of human and murine T11 (CD2) genes. *Proc. Natl. Acad. Sci. USA.* 85:1615.
  16. Lang, G., D. Wotton, M. J. Owen, W. A. Sewell, M. H. Brown, D. Y. Mason, M. J. Crumpton, and D. Kioussis. 1988. The structure of the human CD2 gene and its expression in transgenic mice. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:1675.
  17. Sewell, W. A., M. H. Brown, M. J. Owen, P. J. Fink, C. A. Kozak, and M. J. Crumpton. 1987. The murine homologue of the T lymphocyte CD2 antigen: molecular cloning, chromosome assignment and cell surface expression. *Eur. J. Immunol.* 17:105.
  18. Williams, A. F., A. N. Barclay, S. J. Clark, D. J. Paterson, and A. C. Willis. 1987. Similarities in sequence and cellular expression between rat CD2 and CD4 antigens. *J. Exp. Med.* 165:368.
  19. Taylor, J. W., J. Ott, and F. Eckstein. 1985. The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothionate-modified DNA. *Nucleic Acids Res.* 13:8764.
  20. Haskins, K., R. Kubo, J. White, M. Pigeon, J. Kappler, and P. Marrack. 1983. The major histocompatibility complex-restricted antigen receptor on T cells. I. Isolation with a monoclonal antibody. *J. Exp. Med.* 157:1149.
  21. Cepko, C. L., B. E. Roberts, and R. C. Mulligan. 1984. Construction and applications of highly transmissible murine retrovirus shuttle vector. *Cell.* 37:1053.
  22. Grynkiewicz, G., M. Poenic, and R. Y. Tsien. 1985. A new generation of  $\text{Ca}^{2+}$  indicators with greatly improved fluorescence properties. *J. Biol. Chem.* 260:3440.
  23. Pantaleo, G., D. Olive, A. Poggi, T. Pozzari, L. Moretta, and A. Moretta. 1987. Antibody-induced modulation of the CD3/T cell receptor complex causes T cell refractoriness by inhibiting the early metabolic steps involved in T cell activation. *J. Exp. Med.* 166:619.
  24. Korman, A. J., J. D. Frantz, J. L. Strominger, and R. C. Mulligan. 1987. Expression of human class II major histocompatibility complex antigen using retrovirus vectors. *Proc. Natl. Acad. Sci. USA.* 84:2150.

25. He, Q., A. D. Beyers, A. N. Barclay, and A. F. Williams. 1988. A role in transmembrane signalling for the cytoplasmic domain of the CD2 T lymphocyte surface antigen. *Cell*. 54:979.
26. Bierer, B. A., A. Peterson, J. C. Gorga, S. H. Herrmann, and S. J. Burakoff. 1988. Synergistic T cell activation via the physiological ligands for CD2 and the T cell receptor. *J. Exp. Med.* 168:1145.
27. Alcover, A., C. Alberini, O. Acuto, L. K. Clayton, C. Transy, G. Spagnoli, P. Moingeon, P. Lopez, and E. L. Reinherz. 1988. Interdependence of CD3-Ti and CD2 activation pathways in human T lymphocytes. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:1973.
28. Leo, O., M. Foo, D. H. Sachs, L. E. Samelson, and J. A. Bluestone. 1987. Identification of a monoclonal antibody specific for a murine T3 polypeptide. *Proc. Natl. Acad. Sci. USA*. 84:1374.
29. Gillis, S., M. Ferm, W. Ou, and K. A. Smith. 1978. T cell growth factor: parameters of production and a quantitative microassay for activity. *J. Immunol.* 120:2077.

## Effects of Changes in gp120-CD4 Binding Affinity on Human Immunodeficiency Virus Type 1 Envelope Glycoprotein Function and Soluble CD4 Sensitivity

MARKUS THALI, UDY OLSHEVSKY, CRAIG FURMAN, DANA GABUZDA, JOHN LI,  
AND JOSEPH SODROSKI\*

*Division of Human Retrovirology, Dana-Farber Cancer Institute, and Department of Pathology,  
Harvard Medical School, Boston, Massachusetts 02115*

Received 15 March 1991/Accepted 14 June 1991

**Mutant gp120 glycoproteins exhibiting a range of affinities for CD4 were tested for ability to form syncytia and to complement an *env*-defective provirus for replication. Surprisingly, gp120 mutants that efficiently induced syncytia and/or complemented virus replication were identified that exhibited marked (up to 50-fold) reductions in CD4-binding ability. Temperature-dependent changes in gp120, which result in a seven- to ninefold increase in affinity for CD4, were shown not to be necessary for subsequent membrane fusion or virus entry events. Mutant glycoproteins demonstrating even relatively small decreases in CD4-binding ability exhibited reduced sensitivity to soluble CD4. The considerable range of CD4-binding affinities tolerated by replication-competent HIV-1 variants has important implications for antiviral strategies directed at the gp120-CD4 interaction.**

Human immunodeficiency virus type 1 (HIV-1) is the etiologic agent of AIDS, which is characterized by a depletion of CD4-positive lymphocytes (3, 16, 17, 27, 35). HIV-1 exhibits a tropism for CD4-positive cells due to a high affinity ( $K_d = 4 \times 10^{-9}$  M) interaction between the CD4 glycoprotein, which serves as the viral receptor, and the HIV-1 gp120 exterior envelope glycoprotein (13, 23, 24, 28, 30, 31). Following CD4 binding, the gp120 glycoprotein and gp41, the transmembrane envelope glycoprotein, mediate the fusion of viral and target cell membranes, which is pH independent and necessary for virus entry (18, 25a, 40). Similar events mediated by the envelope glycoproteins expressed on the surface of an infected cell result in fusion of infected cells with CD4-positive cells to form syncytia (18, 25a, 29, 39).

The gp120 molecule binds to a protruding ridge of the CD4 glycoprotein analogous to the CDR2 loop of immunoglobulins (1, 2, 6-8, 21, 26, 32, 34, 36, 42). The CD4-binding site on gp120 is discontinuous, and changes in specific residues located in the second, third, and fourth conserved gp120 regions affect CD4 binding without grossly disrupting the gp120 conformation (10, 11, 18, 25a, 33). Here, we examine the effects of changes in gp120-CD4 affinity on envelope glycoprotein functions involved in syncytium formation and virus entry. We also examine the effect of changes in CD4-binding affinity on the sensitivity of gp120 variants to soluble CD4, which has been shown to exhibit antiviral activity in vitro (4, 14, 15, 19, 38, 41).

### MATERIALS AND METHODS

**Envelope glycoprotein expression and CD4-binding assays.** For the binding assays, COS-1 cells were transfected with pSVIIIenv DNA expressing wild-type or mutant envelope glycoproteins as previously described (12, 18, 33). For initial experiments examining the effect of temperature on binding of the wild-type gp120, supernatants of transfected cells that were radiolabeled with [ $^{35}$ S]cysteine were incubated for

various lengths of time with  $5 \times 10^7$  SupT1 lymphocytes at either 4 or 37°C. The cells were washed once with phosphate-buffered saline (PBS), lysed, and then precipitated with an excess of a mixture of sera derived from AIDS patients (33). The unbound fraction of the supernatants was also used for immunoprecipitation. The amount of gp120 precipitated from both bound and unbound fractions was measured by densitometry of sodium dodecyl sulfate-polyacrylamide gels. Previous studies indicated that the ratio of bound to free gp120 did not vary under these conditions over a greater than 20-fold range of wild-type gp120 concentrations. CD4-binding abilities of the mutant glycoproteins were assessed as described above, by using a 90-min incubation at either 4 or 37°C with SupT1 lymphocytes.

**Syncytium-forming ability of envelope glycoproteins.** To assess the syncytium-forming ability of mutant envelope glycoproteins,  $5 \times 10^5$  COS-1 cells in 100-mm<sup>2</sup> dishes were transfected with 10  $\mu$ g of the envelope expressor plasmid (pSVIIIenv). Forty-eight hours after transfection, the cells were rinsed twice in PBS and incubated with 50 mM EDTA, pH 7.5, at 37°C for approximately 40 min. Cells were removed from the plate by trituration, centrifuged, washed in 5 ml of PBS, and suspended in 5 ml of Dulbecco modified Eagle medium-10% fetal calf serum (DMEM-10% FCS). To these cells was added  $2 \times 10^7$  SupT1 lymphocytes suspended in 5 ml of DMEM-10% FCS. The cells were transferred to a T25 flask and returned to 37°C in a 5% CO<sub>2</sub> incubator, and syncytia were scored in 6 to 8 h.

**Replication complementation assays.** Complementation of the single-step replication of the HXBΔenvCAT provirus into different lymphocytes was performed as previously described (18). Equivalent amounts of reverse transcriptase activity representing recombinant virions in COS-1 supernatants were added to Jurkat, H9, or SupT1 lymphocytes, and chloramphenicol acetyltransferase (CAT) activity was measured at 72 h following infection.

**Soluble CD4 inhibition of syncytium formation and virus replication.** To examine soluble CD4 inhibition of syncytium formation, transfected COS-1 cells detached from tissue

\* Corresponding author.



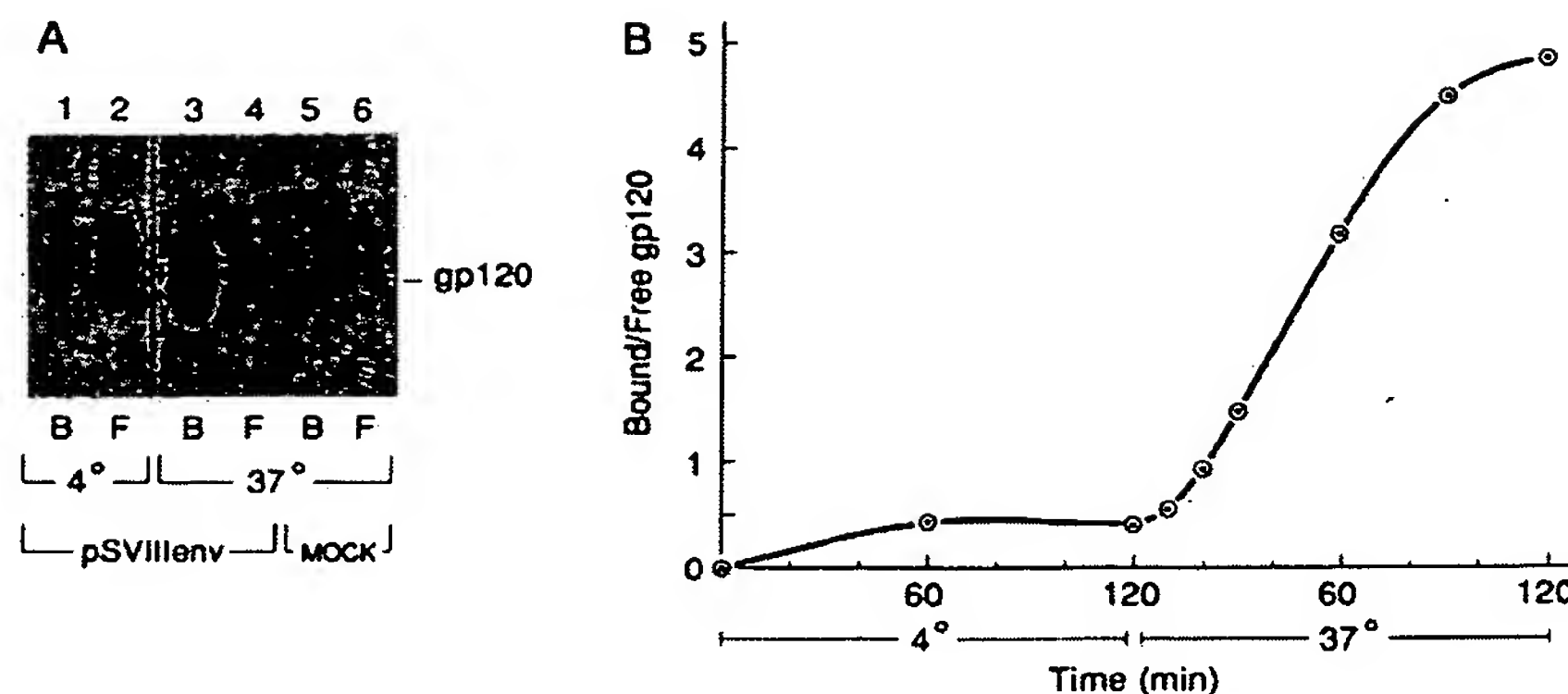


FIG. 1. Effects of temperature on relative CD4 binding. (A) Autoradiogram of wild-type gp120 glycoprotein (lanes 1 through 4) or supernatants from mock-transfected COS-1 cells (lanes 5 and 6) either bound (lanes B: 1, 3, and 5) or unbound (lanes F: 2, 4, and 6) to SupT1 lymphocytes. Experiments were conducted at either 4°C (lanes 1 and 2) or 37°C (lanes 3 through 6). (B) Graph showing the ratio of the amount of wild-type gp120 glycoprotein bound to CD4-positive SupT1 lymphocytes to the amount of unbound gp120 for various times of incubation at either 4 or 37°C. The datum points represent the mean values of data from two independent experiments.

culture dishes as described above were incubated with different concentrations of soluble CD4 (American BioTechnologies, Inc.) in 2.5 ml of DMEM-10% FCS for 1 h at 37°C. To this mixture was added  $2 \times 10^7$  SupT1 lymphocytes suspended in 1 ml of DMEM-10% FCS. Cells were transferred to a T25 flask, and syncytia were scored in 8 h.

For inhibition of virus replication, recombinant HXBΔenv CAT virus produced in COS-1 cells (18) was incubated with different concentrations of soluble CD4 for 1 h at 37°C. The virus preparations were then incubated with Jurkat lymphocytes, and CAT activity was measured 48 h later as previously described (18).

## RESULTS

**Temperature effects on CD4 binding.** The effect of temperature on the ability of radiolabeled soluble gp120 derived from the HXBc2 HIV-1 virus to bind to CD4-positive SupT1 lymphocytes is shown in Fig. 1. Binding equilibrium was achieved by approximately 2 h at either 4 or 37°C. A seven- to ninefold increase in the ratio of bound to free gp120 was observed at 37°C relative to that observed at 4°C. This observed increase at 37°C was only marginally affected by fixing the SupT1 lymphocytes in methanol prior to use (data not shown).

**Effect of changes in gp120 on CD4 binding at different temperatures.** To examine whether changes in the HIV-1 gp120 glycoprotein previously shown to alter its affinity for CD4 (33) affected the temperature-dependent component of the binding process, the binding of gp120 mutants to SupT1 lymphocytes was examined at both 4 and 37°C, using a 90-minute binding assay. The results presented in Table 1 and Fig. 2 demonstrate that although several of the mutants exhibited CD4-binding abilities comparable with or even better than that of the wild-type envelope glycoprotein at 4°C, significant increases in binding at 37°C were not observed for these mutants (e.g., 257 T/R, 368 D/T, 368 D/P, 368 D/E, 370 E/Q, 430 V/S, 457 D/R, and 457 D/A). The relative CD4-binding abilities of the mutants at 37°C, from a 90-min assay, were similar to those previously reported by using a 60-min binding assay (33).

**Function of the envelope glycoprotein mutants.** The ability of the envelope glycoprotein mutants to induce the forma-

tion of syncytia was examined (Table 1). For all of the mutants exhibiting less than one percent of wild-type CD4-binding ability, syncytium-forming ability was undetectable (e.g., 368 D/K, 368 D/R, 370 E/R, 427 W/S, and 427 W/V). Several of the mutants that exhibited smaller decreases in CD4 binding were capable of forming syncytia, although this occurred at a lower efficiency than that observed for the wild-type glycoproteins (e.g., 257 T/R, 368 D/T, 368 D/E, 368 D/N, 370 E/D, 430 V/S, and 457 D/A).

The ability of the mutant envelope glycoproteins to complement the entry of an *env*-defective virus into Jurkat, SupT1, and H9 lymphocytes was assessed. For the mutants exhibiting less than 1% of the wild-type CD4-binding ability, the ability to complement virus entry was less than 10% that of the wild-type envelope glycoprotein in all the target cell types examined (e.g., 368 D/K, 368 D/R, 370 E/R, and 427 W/V). The other mutants, which exhibited CD4-binding ability at 37°C of greater than 1% of the wild-type value, were able to complement virus entry into the target cells examined to various degrees. The 368 D/T, 368 D/E, and 370 E/Q mutants all allowed efficient entry into Jurkat and SupT1 lymphocytes despite approximately 3-, 20-, and 90-fold reductions, respectively, in relative CD4-binding ability. Several of the mutants that did not exhibit temperature-related affinity increases were capable of complementing virus entry into some of the target cell types (257 T/R, 368 D/T, 368 D/P, 368 D/E, 370 E/Q, 430 V/S, and 457 D/A). For these and for other mutants retaining some degree of replication competence, the efficiency of entry into the various target cells utilized in these studies exhibited the pattern Jurkat  $\geq$  SupT1  $\geq$  H9 lymphocytes.

**Soluble CD4 sensitivity of mutant envelope glycoproteins.** The ability of syncytium induction and of virus replication to be inhibited by soluble CD4 was assessed for those mutant glycoproteins competent for these functions (Table 2 and Fig. 3 and 4). Some mutants (120/121 VK/LE, 269 E/L, and 485 K/V) that exhibited mild (less than twofold) reductions in CD4-binding ability were inhibited by soluble CD4 to the same extent as was the wild-type glycoprotein. The functional mutants that exhibited greater than twofold decreases in relative CD4-binding ability were less sensitive than the wild-type glycoprotein to soluble CD4 inhibition of syncy-

TABLE 1. Phenotypes of gp120 mutants<sup>a</sup>

Mutant	Relative CD4 binding at (fold increase) <sup>b</sup>		Syncytium formation	Replication complementation <sup>c</sup> in cell line		
	4°C	37°C		Jurkat	SupT1	H9
Wild type	0.14	1.00 (7.1)	100	100	100	100
257 T/R	0.12	0.06 (0.50)	5	40	25	3
368 D/T	0.31	0.30 (0.97)	16	32	23	7
368 D/P	0.18	0.10 (0.56)	<1	18	19	2
368 D/E	0.05	0.05 (1.00)	51	70	62	22
368 D/N	ND	0.22 (ND)	38	14	11	5
368 D/K	ND	<0.005 (ND)	<1	7	9	2
368 D/R	ND	<0.0004 (ND)	<1	2	2	0
370 E/D	0.14	0.42 (3.0)	73	87	81	50
370 E/Q	0.06	0.011 (0.18)	<1	33	33	3
370 E/R	ND	<0.003 (ND)	<1	1	3	1
427 W/S	ND	<0.012 (ND)	<1	ND	ND	ND
427 W/V	<0.006	<0.006 (ND)	<1	1	ND	1
430 V/S	0.41	0.65 (1.6)	80	100	ND	ND
457 D/R	0.06	0.10 (1.67)	<1	12	9	3
457 D/A	0.08	0.10 (1.25)	47	36	22	6

<sup>a</sup> Data reported represent the means of at least two independent experiments.

<sup>b</sup> The relative CD4-binding ability, determined by using a 90-min binding assay as described in Materials and Methods, is shown. The relative CD4-binding ability was calculated according to the following formula:

$$\text{relative CD4 binding}_{\text{tempX}} = \left[ \frac{(\text{bound mutant gp120})_{\text{tempX}}}{(\text{free mutant gp120})_{\text{tempX}}} \right] \times \left[ \frac{(\text{free wild-type gp120})_{37^\circ\text{C}}}{(\text{bound wild-type gp120})_{37^\circ\text{C}}} \right]$$

The fold increase shown in parentheses represents the ratio of the relative CD4-binding ability of a given mutant at 37°C to that observed at 4°C.

<sup>c</sup> Syncytium formation and replication complementation are represented as the percentage of wild-type values observed for a given target cell. The percent activity observed for a control plasmid not expressing the HIV-1 envelope glycoproteins was less than 1% for syncytium formation and less than 2% for replication complementation in all three target cell lines.

tium formation or virus entry (e.g., 257 T/R, 368 D/T, 368 D/E, 368 D/N, 370 E/D, 370 E/Q, 457 D/A, and 457 D/N). The 430 V/S and 457 D/E mutants exhibited soluble CD4 sensitivities intermediate between that of the wild-type glycoproteins and those of mutants with larger decreases in CD4-binding ability. Mutations affecting the threonine 257 or aspartic acid 457 residues (257 T/A, 457 D/N) that did not decrease CD4-binding ability did not result in soluble CD4 resistance, whereas other changes in these residues (257 T/R, 457 D/A, 457 D/G, and 457 D/E) yielded mutants that exhibited decreased CD4 binding and resistance to soluble CD4.

## DISCUSSION

The interaction of gp120 with CD4 was reduced to undetectable levels by some changes in gp120 residues 368, 370,

or 427. These changes dramatically reduced the ability of the envelope glycoproteins to form syncytia or to complement an *env*-defective virus for replication, consistent with the major contribution of CD4 binding to these processes.

While the gp120 changes that reduced CD4-binding ability by 100-fold or more uniformly abrogated syncytium-forming and/or replicative functions, amino acid changes resulting in less dramatic effects on CD4 binding exerted more variable effects on these envelope glycoprotein functions. The ability of the 368 D/E, 370 E/Q, 370 E/D, and 457 D/A mutants to complement the replication of the *env*-defective virus and the ability of the 368 D/T, 368 D/E, 368 D/N, 370 E/D, 430 V/S, and 457 D/A mutants to form syncytia indicate that even 20- to 50-fold decreases in CD4-binding ability are not necessarily accompanied by abrogation of these functions. These results are consistent with reports that replication-competent HIV-1 isolates can exhibit a broad range of

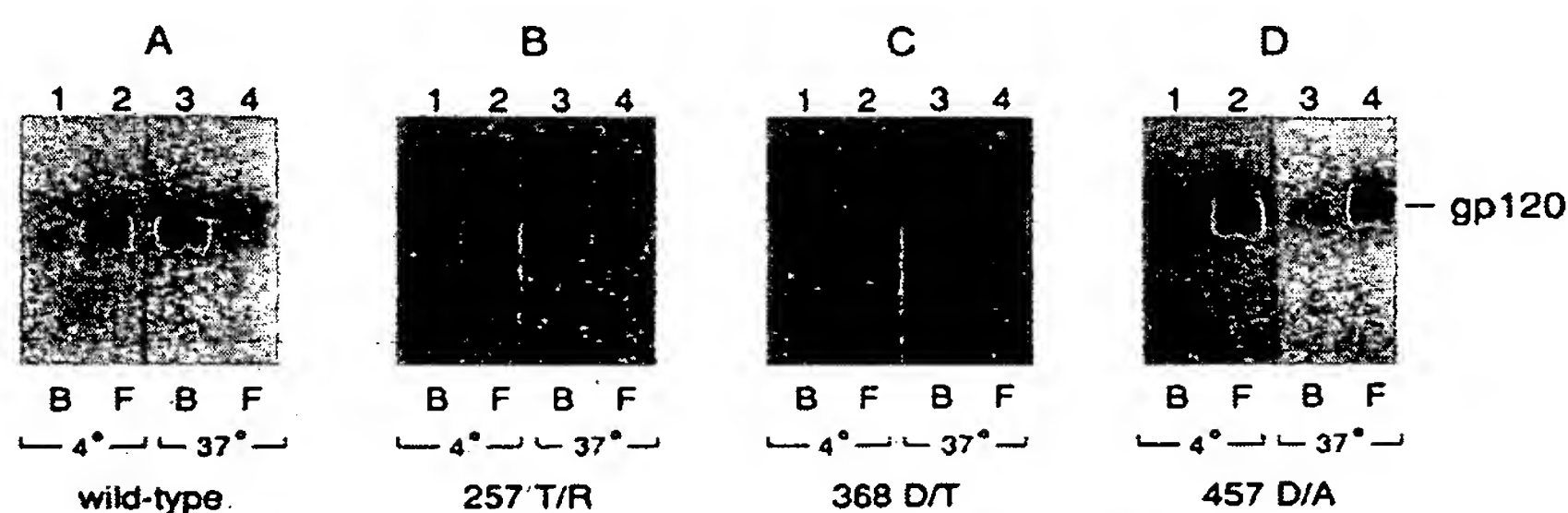


FIG. 2. Effects of amino acid changes on gp120 binding to SupT1 lymphocytes at either 4 or 37°C. The gp120 glycoprotein either bound (lanes B: 1 and 3) or unbound (lanes F: 2 and 4) following a 90-min incubation with SupT1 lymphocytes at either 4°C (lanes 1 and 2) or 37°C (lanes 3 and 4) is shown. The experiments utilized either wild-type gp120 (A) or the 257 T/R, 368 D/T, or 457 D/A mutants (B, C, and D, respectively).

TABLE 2. Soluble CD4 inhibition of envelope glycoprotein function

Mutant	Relative CD4-binding ability <sup>a</sup>	ID <sub>50</sub> syncytium formation (μg/ml) <sup>b</sup>	ID <sub>50</sub> virus replication (μg/ml) <sup>b</sup>
Wild type	1.00	1.5	0.4
113 D/A	1.1	1.5	ND
120/121 VK/LE	0.51	0.9	0.4
257 T/R	0.16	>20	>20
257 T/A	1.1	1.7	0.6
269 E/L	0.61	1.1	0.4
368 D/T	0.33	>20	1.3
368 D/E	0.09	>20	7.5
368 D/N	0.19	>20	ND
370 E/D	0.45	>20	2.5
370 E/Q	0.018	ND	>20
430 V/S	0.39	5.1	1.3
457 D/A	0.09	>20	>20
457 D/N	1.1	1.5	0.4
457 D/G	0.23	ND	>5
457 D/E	0.38	6.5	1.2
485 K/V	0.79	1.0	ND

<sup>a</sup> Relative CD4-binding abilities are derived from experiments in which a 60-min binding assay was used (33).

<sup>b</sup> Values shown are the averages of the results of three independent experiments. Values represent the concentration of soluble CD4 required to inhibit syncytium formation or virus entry by 50%. ND, not determined.

CD4-binding abilities (20). The decreases in syncytium-forming or replicative abilities for some of the mutants (e.g., 257 T/R, 368 D/P, and 457 D/R) that exhibit less than 50-fold reductions in CD4-binding ability are likely to result from the effect of the changes on envelope glycoprotein functions other than receptor binding.

For some of the mutants (257 T/R, 368 D/P, 370 E/Q, and 457 D/R), syncytium formation is affected more dramatically than is the ability to complement virus replication in the same target cells. A similar result has been seen for amino acid changes in the HIV-1 gp41 amino terminus that affect the membrane fusion process (25). These results probably reflect the requirement for a greater number of successful

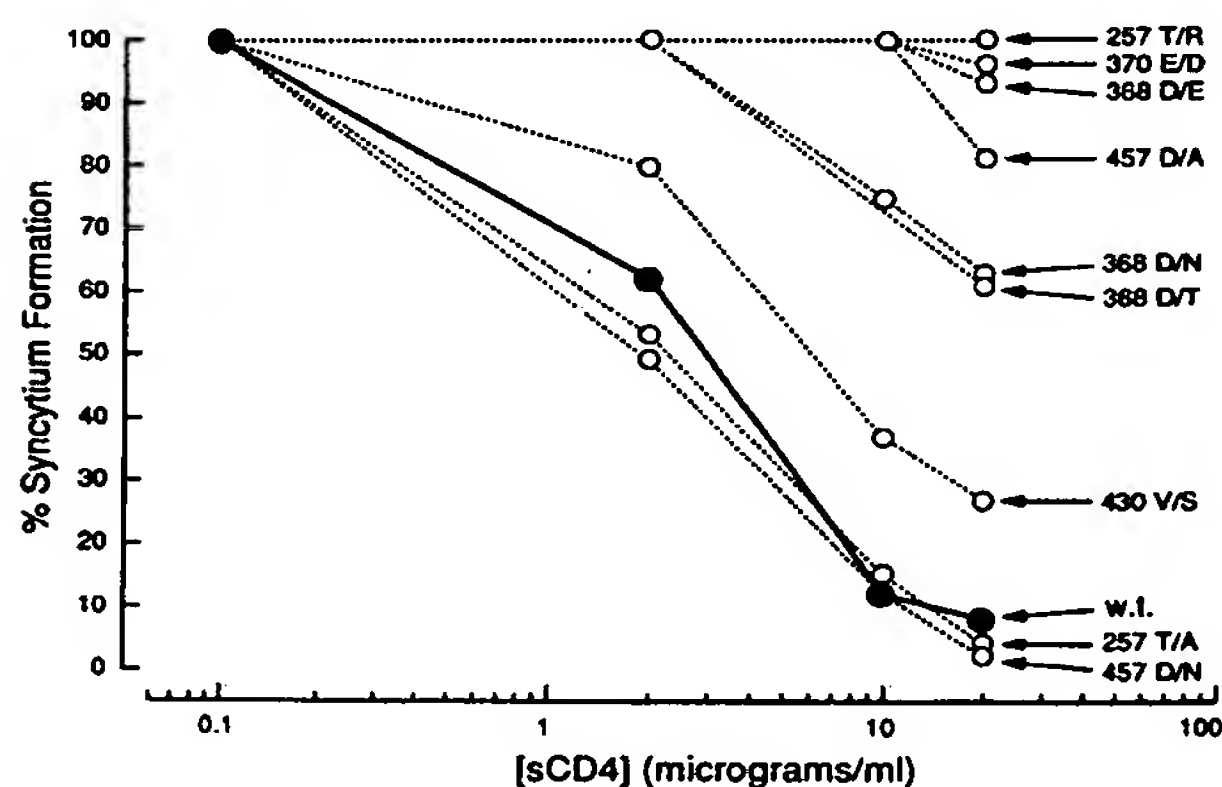


FIG. 3. Effects of amino acid changes in gp120 on inhibition of syncytium formation by soluble CD4. The percentage of the number of syncytia scored in the absence of soluble CD4 is shown for increasing concentrations of soluble CD4. Values are shown for a typical experiment with the wild-type (w.t.) or mutant envelope glycoproteins.

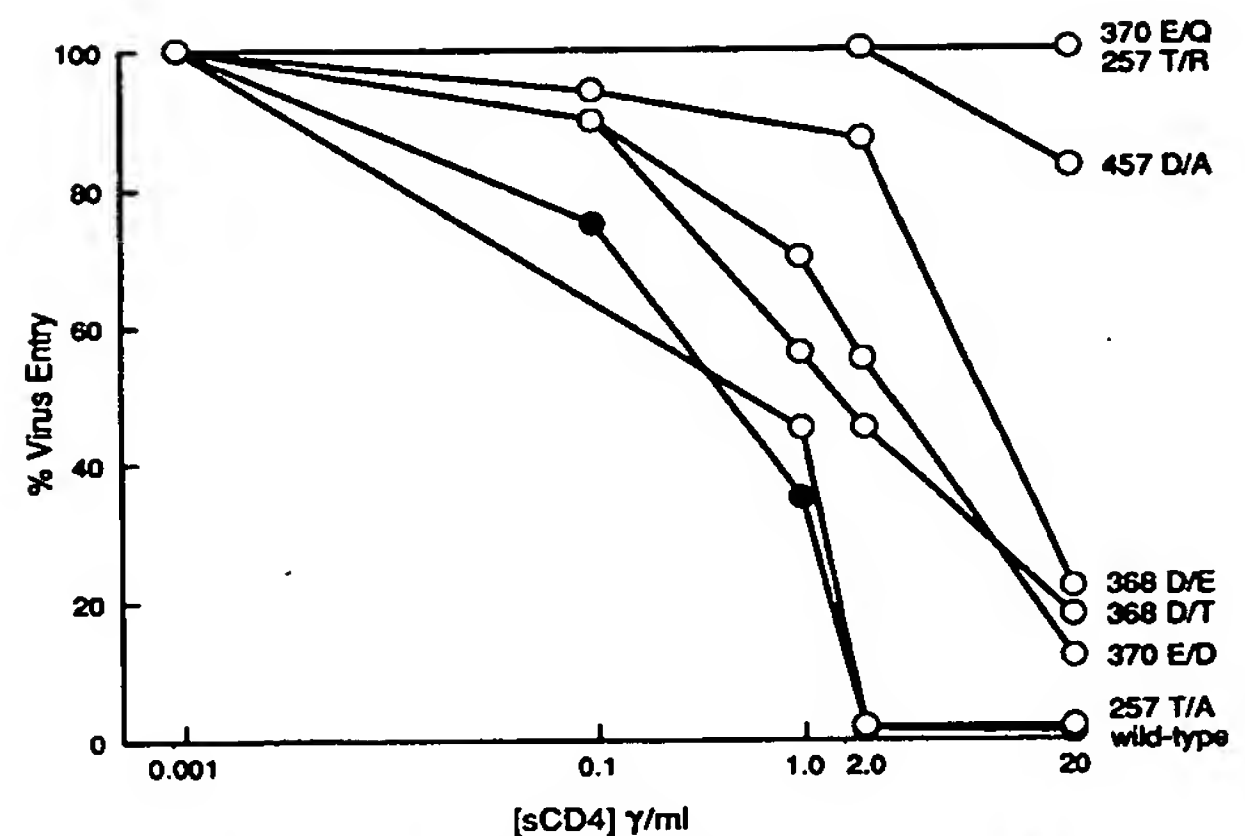


FIG. 4. Effects of gp120 amino acid changes on sensitivity to soluble CD4 inhibition of replication complementation. The effect of different soluble CD4 concentrations on the CAT activity transferred to Jurkat lymphocytes by recombinant virions carrying the wild-type or mutant envelope glycoproteins is shown. Each value represents the percentage of CAT activity observed for a given mutant in the presence of soluble CD4 relative to the activity observed for the mutant in the absence of soluble CD4.

envelope glycoprotein-receptor interactions in the process of cell-cell fusion than in the process of virus entry.

Decreases in CD4-binding ability exerted greater effects on virus entry into some target cells than they did in other cell types. For all of the mutants examined, however, the relative efficiency of entry into various target cells followed the pattern Jurkat  $\geq$  SupT1  $\geq$  H9 lymphocytes. Thus, the efficiency of entry of the mutant virus into these cell types appears to be governed by quantitative factors, one of which may be the level of target cell CD4 expression.

At 4°C, the HIV-1 gp120 glycoprotein bound to CD4-positive cells with an affinity that was approximately seven- to ninefold lower than is seen at 37°C. Following a shift to 37°C, a gradual increase in the association of the wild-type gp120 glycoprotein with CD4-positive cells was observed. This phenomenon cannot be simply explained by endocytosis of cell surface gp120-CD4 complexes, since methanol fixation of the target cells did not eliminate the observed increase. If the effect of low temperature is to limit conformational changes in gp120 or CD4, the results would suggest that induced changes in gp120 and/or CD4 contribute to a higher affinity interaction at 37°C. Mutations affecting several gp120 residues (Thr-257, Asp-368, Glu-370, or Asp-457) decreased the efficiency of these induced changes, suggesting that multiple elements in the protein-protein interface contribute to the affinity increase. The gp120-CD4 interaction, as is true for binding of some antibodies to their antigens (5, 9, 37), might be more appropriately described by an induced fit rather than by a static "lock-and-key" model. Several envelope glycoprotein mutants that did not undergo this temperature-dependent change exhibited significant levels of syncytium-forming or replicative ability, indicating that this conformational alteration is not a necessary step in the membrane fusion process.

Relatively small decreases in CD4-binding ability were associated with envelope glycoprotein mutants resistant to inhibition by soluble CD4. A good correlation was observed between mutants that exhibited a lower affinity for CD4 and



resistance to soluble CD4. This correlation was also observed for mutations affecting the same gp120 amino acid but resulting in different effects on CD4 binding. For example, the 257 T/R, 457 D/A, 457 D/G, and 457 D/E mutants exhibited decreased CD4 binding and were resistant to soluble CD4, whereas the 257 T/A and 457 D/N mutants exhibited approximately wild-type CD4-binding abilities and remained sensitive to soluble CD4. While these results do not rule out other mechanisms for the generation of soluble CD4-resistant HIV-1 variants, they strongly imply that changes in affinity for CD4 represent one such mechanism. Since fresh patient isolates of HIV-1 are often less sensitive to soluble CD4 than are HIV-1 isolates passaged extensively in T-lymphocyte cell lines (12a), some of these primary isolates might exhibit small but consequential decreases in binding affinity for CD4. The sixfold-lower CD4-binding affinity reported for the primary monocyte-propagated Ba-L isolate compared with the T-cell-passaged BH10 isolate might contribute to soluble CD4 resistance (20). HIV-2, which is less sensitive to soluble CD4 than is HIV-1, has been reported to have a 25-fold-lower affinity for CD4 (32a). Such changes in affinity apparently result in major effects on soluble CD4 sensitivity, while exerting only minor effects on virus entry or syncytium formation. This difference may result from the higher cooperativity of envelope glycoprotein-receptor interactions during the latter processes compared with the interaction of soluble CD4 and the envelope glycoproteins. Similar results have recently been reported for the inhibition of polio virus replication by soluble receptor molecules, in which replication-competent viruses resistant to soluble receptor resulted from altered receptor affinity (22).

For some of the mutants exhibiting small changes in relative CD4-binding ability (e.g., 368 D/T and 370 E/D), the 50% inhibitory dose ( $ID_{50}$ ) for soluble CD4 inhibition of syncytium formation is affected more than the corresponding  $ID_{50}$  for virus replication. Soluble CD4 has been reported to act both as a competitive inhibitor and as an irreversible inhibitor, the latter activity resulting from dissociation of the gp120 and gp41 subunits (22a, 32b). Since the soluble CD4-induced subunit dissociation is more efficient for envelope glycoproteins on the virions than on the cell surface (32b), the observations for the mutants could be accounted for by differences in the contributions of reversible and irreversible inhibitory activities of soluble CD4 in the syncytium formation or virus entry process.

Our results have implications for antiviral therapies directed at the gp120-CD4 interaction. The net result of either a reduction in gp120-CD4 affinity or the presence of a competitive inhibitor of the gp120-CD4 interaction is to decrease the ratio of  $gp120_{bound}/gp120_{free}$ . One would expect that the effects of an identical reduction in this ratio on virus replication would be less for a competitive inhibitor than for a mutant exhibiting decreased affinity, since in the former case, the bound gp120 retains wild-type function. Thus, our results imply that competitive inhibitors of the gp120-CD4 interaction will need to be extremely effective to significantly abrogate virus replication, especially in target cells expressing high levels of CD4.

#### ACKNOWLEDGMENTS

We thank Robert Gallo, Flossie Wong-Staal, Max Essex, and Bruce Walker for reagents; Ginny Nixon for manuscript preparation; and Amy Emmert for artwork. We thank American BioTechnologies, Inc., for supplying soluble CD4. Markus Thali was supported by the Swiss National Science Foundation. Udy Olshevsky

performed this work while on sabbatical from the Israel Institute for Biological Research, Nes-Ziona, Israel.

This work was supported by grants from the Leukemia Society of America, the Aaron Diamond Foundation, and the National Institutes of Health (AI 24755).

#### REFERENCES

1. Arthos, J., K. C. Deen, M. Chalkin, J. Fornwald, G. Sathe, Q. Sattentau, P. Clapham, R. Weiss, J. McDougal, C. Pietropaolo, R. Axel, A. Truneh, P. Maddon, and R. Sweet. 1989. Identification of the residues in human CD4 critical for binding of HIV. *Cell* 57:469-481.
2. Ashkenazi, A., L. Presta, S. Marsters, T. Camerato, K. Rosenthal, B. Fendly, and D. Capon. 1990. Mapping the CD4 binding site for HIV by alanine-scanning mutagenesis. *Proc. Natl. Acad. Sci. USA* 87:7150-7154.
3. Barre-Sinoussi, F., J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Danguet, C. Axler-Blin, F. Vezinet-Brun, C. Rouzioux, W. Rozenbaum, and L. Montagnier. 1983. Isolation of a T-lymphocyte retrovirus from a patient at risk for acquired immunodeficiency syndrome (AIDS). *Science* 220:868-871.
4. Berger, E., T. Fuerst, and B. Moss. 1988. A soluble recombinant polypeptide comprising the amino-terminal half of the extracellular region of the CD4 molecule contains an active binding site for human immunodeficiency virus. *Proc. Natl. Acad. Sci. USA* 85:2357-2361.
5. Bhat, T., G. Bentley, T. Fischmann, G. Boulot, and R. Poljak. 1990. Small rearrangements in structures of Fv and Fab fragments of antibody D1.3 on antigen binding. *Nature (London)* 347:483-485.
6. Bowman, M., K. MacFerrin, S. Schreiber, and S. Burakoff. 1990. Identification and structural analysis of residues in the V1 region of CD4 involved in interaction with human immunodeficiency virus envelope glycoprotein gp120 and class II major histocompatibility complex molecules. *Proc. Natl. Acad. Sci. USA* 87:9052-9056.
7. Brodsky, M., M. Wharton, R. Myers, and D. Littman. 1990. Analysis of the site in CD4 that binds to the HIV envelope glycoprotein. *J. Immunol.* 144:3078-3086.
8. Clayton, L., R. Hussey, R. Steinbrich, H. Ramachandran, Y. Husain, and E. Reinherz. 1988. Substitution of murine for human CD4 residues identifies amino acids critical for HIV gp120 binding. *Nature (London)* 335:363-366.
9. Colman, P., W. Laver, J. Varghese, A. Baker, P. Tullock, G. Air, and R. Webster. 1987. Three dimensional structure of a complex of antibody with influenza virus neuraminidase. *Nature (London)* 326:358-363.
10. Cordonnier, A., L. Montagnier, and M. Emerman. 1989. Single amino acid changes in HIV envelope affect viral tropism and receptor binding. *Nature (London)* 340:571-574.
11. Cordonnier, A., Y. Riviere, L. Montagnier, and M. Emerman. 1989. Effects of mutations in hyperconserved regions of the extracellular glycoprotein of human immunodeficiency virus type 1 on receptor binding. *J. Virol.* 63:4464-4468.
12. Cullen, B. R. 1987. Use of eukaryotic expression technology in the functional analysis of cloned genes. *Methods Enzymol.* 152:664-703.
- 12a. Daar, E., X. L. Li, T. Moudgil, and D. Ho. 1990. High concentrations of recombinant soluble CD4 are required to neutralize primary human immunodeficiency virus type 1 isolates. *Proc. Natl. Acad. Sci. USA* 87:6574-6578.
13. Dalgleish, A. G., P. C. L. Beverley, P. R. Clapham, D. H. Crawford, M. F. Greaves, and R. A. Weiss. 1984. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature (London)* 312:763-767.
14. Deen, K., J. S. McDougal, R. Inacker, G. Folena-Wasserman, J. Arthos, J. Rosenberg, P. Maddon, R. Axel, and R. Sweet. 1988. A soluble form of CD4 (T4) protein inhibits AIDS virus infection. *Nature (London)* 331:82-84.
15. Fisher, R., J. Bertonis, W. Meier, V. Johnson, D. Costopoulos, T. Liu, R. Tizard, B. Walker, M. Hirsch, R. Schooley, and R. Flavell. 1988. HIV infection is blocked *in vitro* by recombi-

- nant soluble CD4. *Nature (London)* 331:76-78.
16. Gallo, R. C., S. Z. Salahuddin, M. Popovic, G. M. Shearer, M. Kaplan, B. F. Hayer, T. J. Palker, R. Redfield, J. Oleske, B. Safai, G. White, P. Foster, and P. D. Markham. 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* 224:500-503.
  17. Gottlieb, M. S. 1986. Immunologic aspects of the acquired immunodeficiency syndrome and male homosexuality. *Med. Clin. North Am.* 70:651-654.
  18. Helseth, E., M. Kowalski, D. Gabuzda, U. Olshevsky, W. Haseltine, and J. Sodroski. 1990. Rapid complementation assays measuring replicative potential of human immunodeficiency virus type 1 envelope glycoprotein mutants. *J. Virol.* 64:2416-2420.
  19. Hussey, R., N. Richardson, M. Kowalski, N. Brown, H. Change, R. Siliciano, T. Dorfman, B. Walker, J. Sodroski, and E. Reinherz. 1988. A soluble CD4 protein selectively inhibits HIV replication and syncytium formation. *Nature (London)* 331:78-81.
  20. Ivey-Hoyle, M., J. Culp, M. Chaikin, B. Helling, T. Matthews, R. Sweet, and M. Rosenberg. 1991. Envelope glycoproteins from biologically diverse isolates of immunodeficiency viruses have widely different affinities for CD4. *Proc. Natl. Acad. Sci. USA* 88:512-516.
  21. Jameson, B., P. Rao, L. Kong, B. Hahn, G. Shaw, L. Hood, and S. Kent. 1988. Location and chemical synthesis of a binding site for HIV-1 on the CD4 protein. *Science* 240:1335-1338.
  22. Kaplan, G., D. Peters, and V. Racaniello. 1990. Poliovirus mutants resistant to neutralization with soluble cell receptors. *Science* 250:1596-1599.
  - 22a. Kirsh, R., T. Hart, H. Ellens, J. Miller, S. Petteway, D. Lambert, B. Leary, and P. Bugelski. 1990. Morphometric analysis of recombinant soluble CD4-mediated release of the envelope glycoprotein gp120 from HIV-1. *AIDS Res. Hum. Retroviruses* 6:1209-1212.
  23. Klatzmann, D., F. Barre-Sinoussi, T. Nugeyre, C. Dauguet, E. Vilmer, C. Griscelli, F. Brun-Vezinet, C. Rouzioux, J. C. Gluckman, J. C. Chermann, and L. Montagnier. 1984. Selective tropism of lymphadenopathy-associated virus (LAV) for helper-inducer T lymphocytes. *Science* 225:59-63.
  24. Klatzmann, D., E. Champagne, S. Chamaret, J. Gruest, D. Guetard, T. Hercend, J. C. Gluckman, and L. Montagnier. 1984. T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature (London)* 312:767-768.
  25. Kowalski, M., L. Bergeron, T. Dorfman, W. Haseltine, and J. Sodroski. 1991. Attenuation of human immunodeficiency virus type 1 cytopathic effect by a mutation affecting the transmembrane envelope glycoprotein. *J. Virol.* 65:281-291.
  - 25a. Kowalski, M., J. Potz, L. Basiripour, T. Dorfman, W. C. Goh, E. Terwilliger, A. Dayton, C. Rosen, W. A. Haseltine, and J. Sodroski. 1987. Functional regions of the human immunodeficiency virus envelope glycoprotein. *Science* 237:1351-1355.
  26. Landau, N. R., M. Warton, and D. R. Littman. 1988. The envelope glycoprotein of the human immunodeficiency virus binds to the immunoglobulin-like domain of CD4. *Nature (London)* 334:159-161.
  27. Lane, M. C., J. L. Depper, W. C. Greene, G. Whalen, T. Waldmann, and A. S. Fauci. 1985. Qualitative analysis of immune function in patients with the acquired immunodeficiency syndrome. *N. Engl. J. Med.* 313:79-84.
  28. Lasky, L. A., G. M. Nakamura, D. H. Smith, C. Fennie, C. Shimasaki, E. Patzer, O. Berman, T. Gregory, and D. J. Capon. 1987. Delineation of a region of the human immunodeficiency virus type 1 gp120 glycoprotein critical for interaction with the CD4 receptor. *Cell* 50:975-985.
  29. Lifson, J., M. Feinberg, G. Reyes, L. Rabin, B. Banapour, S. Chakrabarti, B. Moss, F. Wong-Staal, K. Steimer, and E. Engleman. 1986. Induction of CD4-dependent cell fusion by the HTLV-III/LAV envelope glycoprotein. *Nature (London)* 323:725-728.
  30. Maddon, P., A. Dalglish, J. S. McDougal, P. Clapham, R. Weiss, and R. Axel. 1986. The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell* 47:333-348.
  31. McDougal, J. S., M. Kennedy, J. Sligh, S. Cort, A. Mowle, and J. Nicholson. 1986. Binding of the HTLV-III/LAV to T4<sup>+</sup> T cells by a complex of the 100 K viral protein and the T4 molecule. *Science* 231:382-385.
  32. Mizukami, T., T. Fuerst, E. Berger, and B. Moss. 1988. Binding region for human immunodeficiency virus (HIV) and epitopes for HIV-blocking monoclonal antibodies of the CD4 molecule defined by site-directed mutagenesis. *Proc. Natl. Acad. Sci. USA* 85:9273-9277.
  - 32a. Moore, J. 1990. Simple methods for monitoring HIV-1 and HIV-2 gp120 binding to soluble CD4 by enzyme-linked immunosorbent assay: HIV-2 has a 25-fold lower affinity than HIV-1 for soluble CD4. *AIDS* 4:297-305.
  - 32b. Moore, J., J. McKeating, R. Weiss, and Q. Sattentau. 1990. Dissociation of gp120 from HIV-1 virions induced by soluble CD4. *Science* 250:1139-1142.
  33. Olshevsky, U., E. Helseth, C. Furman, J. Li, W. Haseltine, and J. Sodroski. 1990. Identification of individual human immunodeficiency virus type 1 gp120 amino acids important for CD4 receptor binding. *J. Virol.* 64:5701-5707.
  34. Peterson, A., and B. Seed. 1988. Genetic analysis of monoclonal antibodies and HIV binding sites on the human lymphocyte antigen CD4. *Cell* 54:65-72.
  35. Popovic, M., M. G. Sarngadharan, E. Read, and R. C. Gallo. 1984. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science* 224:497-500.
  36. Ryu, S.-E., P. Kwong, A. Truneh, T. Porter, J. Arthos, M. Rosenberg, X. Dai, N. Xuong, R. Axel, R. Sweet, and W. Hendrickson. 1990. Crystal structure of an HIV-binding recombinant fragment of human CD4. *Nature (London)* 348:419-425.
  37. Sheriff, S., E. Silverton, E. Padlan, G. Cohen, S. Smith-Gill, B. Finzel, and D. Davies. 1987. Three-dimensional structure of an antibody-antigen complex. *Proc. Natl. Acad. Sci. USA* 84:8075-8079.
  38. Smith, D., R. Byrn, S. Marsters, T. Gregory, J. Groopman, and D. Capon. 1987. Blocking of HIV-1 infectivity by a soluble secreted form of the CD4 antigen. *Science* 238:1704-1707.
  39. Sodroski, J., W. C. Goh, C. A. Rosen, K. Campbell, and W. Haseltine. 1986. Role of the HTLV-III envelope in syncytium formation and cytopathicity. *Nature (London)* 321:412-417.
  40. Stein, B. S., S. D. Gouda, J. D. Lifson, R. C. Penhallow, K. G. Bensch, and E. G. Engleman. 1987. pH-independent HIV entry into CD4-positive T cells via virus envelope fusion to the plasma membrane. *Cell* 49:659-668.
  41. Trauneker, A., W. Luke, and K. Karjalainen. 1988. Soluble CD4 molecules neutralize human immunodeficiency virus type 1. *Nature (London)* 331:84-86.
  42. Wang, J., Y. Yan, T. Garrett, J. Liu, D. Rodgers, R. Garlick, G. Tarr, Y. Husain, E. Reinherz, and S. Harrison. 1990. Atomic structure of a fragment of human CD4 containing two immunoglobulin-like domains. *Nature (London)* 348:411-418.



## Vascular Endothelial Growth Factor Induces Expression of Connective Tissue Growth Factor via KDR, Flt1, and Phosphatidylinositol 3-Kinase-Akt-dependent Pathways in Retinal Vascular Cells\*

Received for publication, July 21, 2000, and in revised form, September 28, 2000  
Published, JBC Papers in Press, October 3, 2000, DOI 10.1074/jbc.M006509200

Kiyoshi Suzuma†§, Keiko Naruse‡, Izumi Suzuma‡, Noriko Takahara‡, Kohjiro Ueki‡, Lloyd P. Aiello†||, and George L. King‡\*\*

From the ‡Research Division and ||Beetham Eye Institute, Joslin Diabetes Center, and the †Department of Ophthalmology, Harvard Medical School, Boston, Massachusetts 02215

Fibroblastic proliferation accompanies many angiogenesis-related retinal and systemic diseases. Since connective tissue growth factor (CTGF) is a potent mitogen for fibrosis, extracellular matrix production, and angiogenesis, we have studied the effects and mechanism by which vascular endothelial growth factor (VEGF) regulates CTGF gene expression in retinal capillary cells. In our study, VEGF increased CTGF mRNA levels in a time- and concentration-dependent manner in bovine retinal endothelial cells and pericytes, without the need of new protein synthesis and without altering mRNA stability. VEGF activated the tyrosine receptor phosphorylation of KDR and Flt1 and increased the binding of phosphatidylinositol 3-kinase (PI3-kinase) p85 subunit to KDR and Flt1, both of which could mediate CTGF gene induction. VEGF-induced CTGF expression was mediated primarily by PI3-kinase activation, whereas PKC and ERK pathways made only minimal contributions. Furthermore, overexpression of constitutive active Akt was sufficient to induce CTGF gene expression, and inhibition of Akt activation by overexpressing dominant negative mutant of Akt abolished the VEGF-induced CTGF expression. These data suggest that VEGF can increase CTGF gene expression in bovine retinal capillary cells via KDR or Flt receptors and the activation of PI3-kinase-Akt pathway independently of PKC or Ras-ERK pathway, possibly inducing the fibrosis observed in retinal neovascular diseases.

Angiogenesis and fibrosis are key components in development, growth, wound healing, and regeneration (1). In addition, these processes commonly occur together in many disease states where neovascularization is believed to initiate the pathological cascade. Some of these diseases are proliferative diabetic retinopathy (2), rheumatoid arthritis (3), and age-related macular degeneration (4). Thus, it is possible that the factors that regulate angiogenesis may also induce factors that stimulate extracellular matrix production and fibrosis. Accordingly,

we have studied the ability of vascular endothelial growth factor (VEGF),<sup>1</sup> an established angiogenic factor, to regulate the expression of connective tissue growth factor (CTGF), a growth factor with known actions on fibroblast proliferation, matrix production, and associated with fibrotic disorders.

VEGF is expressed as a family of peptides of 121, 145, 165, 189, and 206 amino acid residues (5). Its expression is induced by hypoxia (6) and is essential in the vasculogenesis process during development (7). Several receptors have been shown to mediate the action of VEGF, and most of them belong to the tyrosine kinase receptor family (8). Upon the binding of VEGF to its receptors, multiple signaling cascades are activated, including the tyrosine phosphorylation of phospholipase C $\gamma$ , elevation of intracellular calcium and diacylglycerol, activation of protein kinase C (PKC), and extracellular signal-regulated kinase (MAPK/ERK) for endothelial cell proliferation (9–12). In addition, VEGF also stimulates activation of phosphatidylinositol (PI) 3-kinase leading to Akt/PKB activation and possibly enhancing endothelial cell survival (13–15). However, in non-endothelial cells such as capillary pericytes that predominantly express Flt1 receptor, the action of VEGF is poorly understood.

Connective tissue growth factor (CTGF), a member of CCN family (CYR61, CTGF, and NOV) (16, 17), is a potent and ubiquitously expressed growth factor that has been shown to play a unique role in fibroblast proliferation, cell adhesion, and the stimulation of extracellular matrix production (18, 19). The 38-kDa protein was originally identified in conditioned medium from human umbilical vein endothelial cells (20), and the expression was shown to be selectively stimulated by transforming growth factor- $\beta$  (TGF- $\beta$ ) in cultured fibroblasts (21). Due to its mitogenic action on fibroblasts and its ability to induce the expression of the extracellular matrix molecules, collagen type I, fibronectin, and integrin  $\alpha_5$  (18), CTGF is supposed to play an important role in connective tissue cell proliferation and extracellular matrix deposition as one of the mediators of TGF- $\beta$

\* This work was supported in part by National Institutes of Health Grants EY5110 (to G. L. K.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Recipient of the Mary K. Iacocca Fellowship.

\*\* To whom correspondence should be addressed: Research Division, Joslin Diabetes Center, One Joslin Place, Boston, MA 02215. Tel.: 617-732-2622; Fax: 617-732-2637; E-mail: George.King@joslin.harvard.edu.

<sup>1</sup> The abbreviations used are: VEGF, vascular endothelial growth factor; CTGF, connective tissue growth factor; Flt, *fms*-like tyrosine kinase; KDR/Flk1, fetal liver kinase/kinase domain-containing receptor; VEGFR, vascular endothelial growth factor receptor; PKB, protein kinase B; MAPK, mitogen-activated protein kinase; ERK, extracellular signal-regulated kinase; kb, kilobase pair; PKC, protein kinase C; PI, phosphatidylinositol; TGF- $\beta$ , transforming growth factor- $\beta$ ; BREC, bovine retinal endothelial cells; PlGF, placenta growth factor; BRPC, bovine retinal pericytes; PCR, polymerase chain reaction; CAAkt, constitutive active Akt; DNakt, dominant negative Akt; DNRas, dominant negative K-Ras; DNERK, dominant negative extracellular signal-regulated kinase; DNPCK $\zeta$ , dominant negative PKC $\zeta$ .



(22). CTGF also seems to be an important player in the pathogenesis of various fibrotic disorders, since it was shown to be overexpressed in scleroderma, keloids, and other fibrotic skin disorders (23), as well as in stromal rich mammary tumors (24), and in advanced atherosclerotic lesions (25). Recently, the integrin  $\alpha_v\beta_3$  has been reported to serve as a receptor on endothelial cells for CTGF-mediated endothelial cell adhesion, migration, and angiogenesis (26, 27).

Besides TGF- $\beta$ , the expression of CTGF is reported to be regulated by dexamethasone in BALB/c 3T3 cells (28), high glucose in human mesangial cells (29), kinin in human embryonic fibroblasts (30), factor VIIa, and thrombin in WI-38 fibroblasts (31), tumor necrosis factor  $\alpha$  in human skin fibroblast (32), and cAMP in bovine endothelial cells (33). Since many of these cytokines are known to induce VEGF, it is possible that increased VEGF expression can regulate the expression of CTGF. In the present study, we have investigated the regulation of CTGF by VEGF in retinal endothelial cells and pericytes via the PI3-kinase and several other signaling pathways.

#### EXPERIMENTAL PROCEDURES

**Materials**—Endothelial cell basal medium was purchased from Clonetics (San Diego, CA). Endothelial cell growth factor was purchased from Roche Molecular Biochemicals. Dulbecco's modified Eagle's medium and fetal bovine serum were obtained from Life Technologies, Inc. VEGF, placenta growth factor (PlGF), TGF- $\beta$ 1, and anti-CTGF antibody were ordered from R & D Systems (Minneapolis, MN). Anti-KDR (Flk1) and anti-Flt1 antibodies were purchased from Santa Cruz Biotechnology (Santa Cruz, CA). Protein A-Sepharose was purchased from Amersham Pharmacia Biotech. Anti-phospho-ERK, anti-ERK, anti-phospho-Akt, and anti-Akt were purchased from New England Biolabs (Beverly, MA). Anti-p85 and anti-phosphotyrosine were purchased from Upstate Biotechnology, Inc. (Lake Placid, NY). Phosphatidylinositol (PI) was purchased from Avanti (Alabaster, AL), and PD98059, wortmannin, and GF 109203X were obtained from Calbiochem. All other materials were ordered from Fisher and Sigma.

**Cell Culture**—Primary cultures of bovine retinal endothelial cells (BREC) and pericytes (BRPC) were isolated by homogenization and a series of filtration steps as described previously (34). BREC were subsequently cultured with endothelial cell basal medium supplemented with 10% plasma-derived horse serum, 50 mg/liter heparin, and 50  $\mu$ g/ml endothelial cell growth factor. BRPC were cultured in Dulbecco's modified Eagle's medium with 5.5 mM glucose and 20% fetal bovine serum. Cells were cultured in 5% CO<sub>2</sub> at 37 °C, and media were changed every 3 days. Cells were characterized for their homogeneity by immunoreactivity with anti-factor VIII antibody for BREC and with monoclonal antibody 3G5 for BRPC (35). Cells remained morphologically unchanged under these conditions, as confirmed by light microscopy. Only cells from passages 2 through 7 were used for the experiments.

**Recombinant Adenoviruses**—cDNA of constitutive active Akt (CAAkt, Gag protein fused to N-terminal of wild type Akt) was constructed as described (36). cDNA of dominant negative Akt (DNAkt, substituted Thr-308 to Ala and Ser-473 to Ala) was constructed as described (37). cDNA of dominant negative K-Ras (DNK-Ras, substituted Ser-17 to Asn) was kindly provided by Dr. Takai (Osaka University) (38). cDNA of dominant negative extracellular signal-regulated kinase (DNERK, substituted Lys-52 to Arg in ATP-binding site) was constructed as described (39). cDNA of  $\Delta$ p85 was kindly provided by Dr. Kasuga (Kobe University) (40). cDNA of PKC $\zeta$  was kindly provided by Dr. Douglas Ways (Lilly). cDNA of dominant negative PKC $\zeta$  (DNPKC $\zeta$ , substituted Lys-273 to Trp in ATP-binding site) was constructed as described (41). The recombinant adenoviruses were constructed by homologous recombination between the parental virus genome and the expression cosmid cassette or shuttle vector as described (42, 43). The adenoviruses were applied at a concentration of  $1 \times 10^8$  plaque-forming units/ml, and adenoviruses with the same parental genome carrying the *lacZ* gene or enhanced green fluorescent protein gene (CLONTECH, Palo Alto, CA) were used as controls. Expression of each recombinant protein was confirmed by Western blot analysis and increased about 10-fold compared with cells infected with the control adenovirus.

**Immunoprecipitation**—Cells were washed three times with cold phosphate-buffered saline and solubilized in 200  $\mu$ l of lysis buffer (1% Triton X-100, 50 mmol/liter HEPES, 10 mmol/liter EDTA, 10 mmol/liter sodium pyrophosphate, 100 mmol/liter sodium fluoride, 1 mmol/liter sodium orthovanadate, 1  $\mu$ g/ml aprotinin, 1  $\mu$ g/ml leupeptin, and 2

mmol/liter phenylmethylsulfonyl fluoride). After centrifugation at 12,000 rpm for 10 min, 1.0 mg of protein was subjected to immunoprecipitation. To clear the protein extract, protein A-Sepharose (20  $\mu$ l of a 50% suspension) was added to the cell lysates, after which they were incubated for 1 h, followed by centrifugation and collection of the supernatant. A specific rabbit anti-KDR or Flt1 antibody was added and rocked at 4 °C for 2 h; 20  $\mu$ l of protein A-Sepharose was then added, and the sample was rocked for another 2 h at 4 °C. For denaturation, protein A-Sepharose antigen-antibody conjugates were separated by centrifugation, washed five times, and boiled for 3 min in Laemmli sample buffer.

**Western Blot Analysis**—Immunoprecipitated proteins or 30  $\mu$ g of total cell lysates were subjected to SDS-gel electrophoresis and electrotransferred to nitrocellulose membrane (Bio-Rad). The membrane was soaked in blocking buffer (phosphate-buffered saline containing 0.1% Tween 20 and 5% bovine serum albumin) for 1 h at room temperature and incubated with primary antibody overnight at 4 °C followed by incubation with horseradish peroxidase-conjugated secondary antibody (Amersham Pharmacia Biotech). Visualization was performed using the enhanced chemiluminescence detection system (ECL, Amersham Pharmacia Biotech) per the manufacturer's instructions.

**PI3-Kinase Assay**—PI3-kinase activities were measured by the *in vitro* phosphorylation of PI (10). Cells were lysed in ice-cold lysis buffer containing 50 mM HEPES, pH 7.5, 137 mM NaCl, 1 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub>, 2 mM Na<sub>3</sub>VO<sub>4</sub>, 10 mM NaF, 2 mM EDTA, 1% Nonidet P-40, 10% glycerol, 1 mM phenylmethylsulfonyl fluoride, 2  $\mu$ g/ml aprotinin, 5  $\mu$ g/ml leupeptin, and 1  $\mu$ g/ml pepstatin. Insoluble material was removed by centrifugation at 15,000  $\times g$  for 10 min at 4 °C. PI3-kinase was immunoprecipitated from aliquots of the supernatant with anti-phosphotyrosine antibodies. After successive washings, the pellets were resuspended in 50  $\mu$ l of 10 mM Tris, pH 7.5, 100 mM NaCl, and 1 mM EDTA. 10  $\mu$ l of 100 mM MgCl<sub>2</sub> and 10  $\mu$ l of PI (2  $\mu$ g/ $\mu$ l) sonicated in 10 mM Tris, pH 7.5, with 1 mM EGTA was added to each pellet. The PI3-kinase reaction was initiated by the addition of 5  $\mu$ l of 0.5 mM ATP containing 30  $\mu$ Ci of [ $\gamma$ -<sup>32</sup>P]ATP. After 10 min at room temperature with constant shaking, the reaction was stopped by the addition of 20  $\mu$ l of 8 N HCl and 160  $\mu$ l of chloroform/methanol (1:1). The samples were centrifuged, and the organic phase was removed and applied to silica gel TLC plates developing in CHCl<sub>3</sub>/CH<sub>3</sub>OH/H<sub>2</sub>O/NH<sub>4</sub>OH (60:47:11:2). The radioactivity in spots was quantified by PhosphorImager (Molecular Dynamics, Sunnyvale, CA).

**Amplification of Human CTGF cDNA Using Reverse Transcriptase-Polymerase Chain Reaction (PCR)**—cDNA templates for PCR were synthesized by reverse transcriptase (First Strand cDNA Synthesis Kit, Amersham Pharmacia Biotech) from human fibroblast according to the method recommended by the manufacturer. A standard PCR was performed (PCR optimizer kit, Invitrogen, Carlsbad, CA) using 5'-AGGG-CCTCTTCTGTGACTTCG-3' (sense primer) and 5'-TCATGCCATGTC-TCCGTACATC-3' (antisense primer) (20). The PCR products were then subcloned into a vector (pCRII, Invitrogen) and sequenced in their entirety, and comparison with the published human sequences revealed complete sequence identity. This cDNA probe was used for hybridization.

**Northern Blot Analysis**—Total RNA was isolated using acid-guanidinium thiocyanate, and Northern blot analysis was performed as described previously (44). Total RNA (20  $\mu$ g) was electrophoresed through 1% formaldehyde-agarose gels and then transferred to a nylon membrane. <sup>32</sup>P-Labeled cDNA probes were generated by use of labeling kits (Megaprime DNA labeling systems, Amersham Pharmacia Biotech). After ultraviolet cross-linking using a UV cross-linker (Stratagene, La Jolla, CA), blots were pre-hybridized, hybridized, and washed in 0.5 $\times$  SSC, 5% SDS at 65 °C with 4 changes over 1 h. All signals were analyzed using a PhosphorImager, and lane loading differences were normalized using the 36B4 cDNA probe (45).

**Analysis of CTGF mRNA Half-life**—CTGF mRNA half-life experiments were carried out using BREC and BRPC. The cells were exposed to vehicle or VEGF (25 ng/ml) for the indicated periods prior to mRNA stability measurements. Transcription was inhibited by the addition of actinomycin D (5  $\mu$ g/ml). For inhibition of protein synthesis, cells were treated with cycloheximide (10  $\mu$ g/ml) for the times indicated.

**Statistical Analysis**—Determinations were performed in triplicate, and all experiments were repeated at least three times. Results are expressed as the mean  $\pm$  S.D., unless otherwise indicated. Statistical analysis employed Student's *t* test or analysis of variance to compare quantitative data populations with normal distributions and equal variance. Data were analyzed using the Mann-Whitney rank sum test or the Kruskal-Wallis test for populations with non-normal distributions or unequal variance. A *p* value of <0.05 was considered statistically significant.

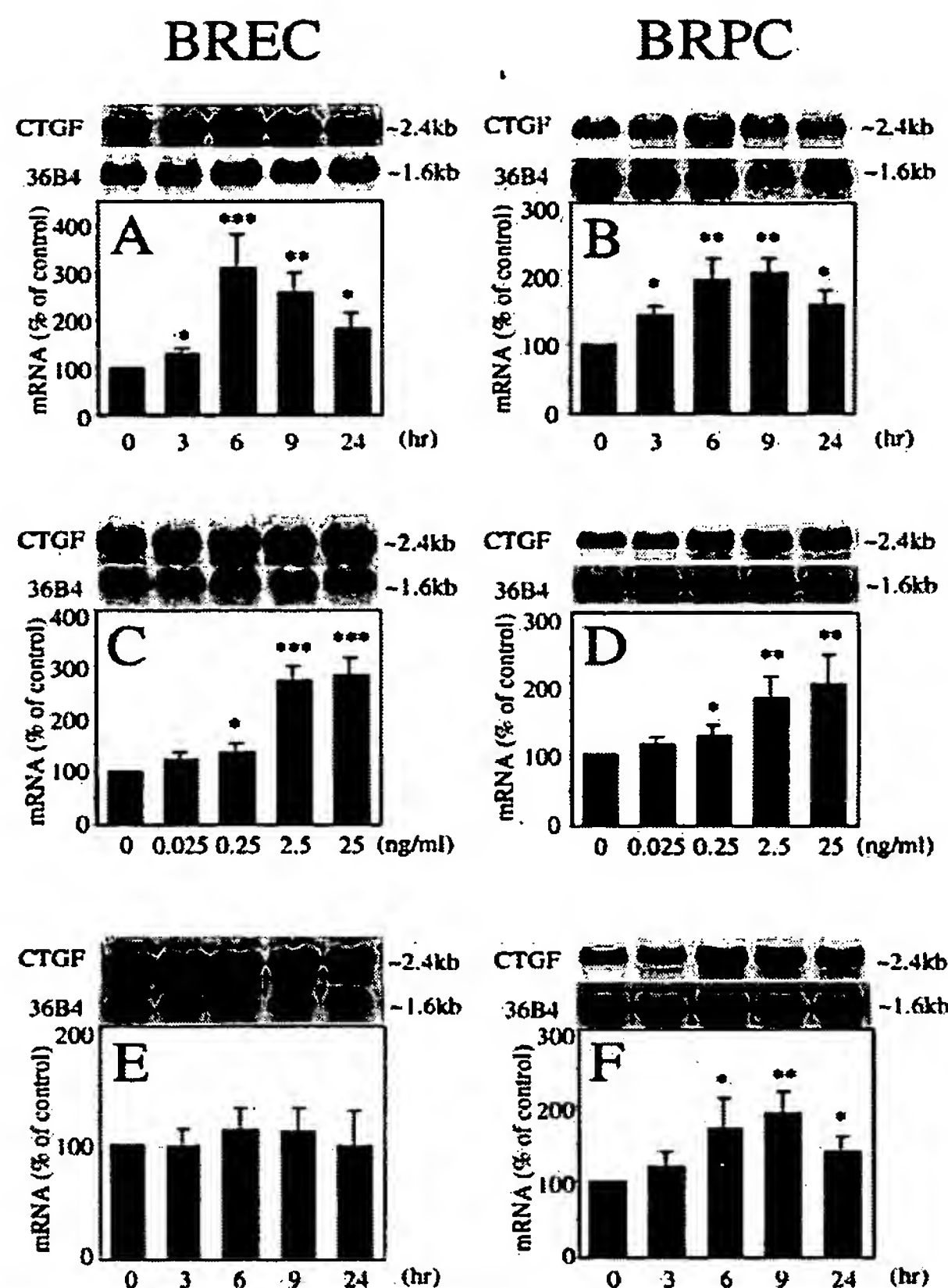


FIG. 1. Stimulation of CTGF mRNA expression by VEGF and PlGF. A (BREC) and B (BRPC), time-dependent induction of CTGF mRNA after stimulation by 25 ng/ml VEGF. C (BREC) and D (BRPC), dose-dependent induction of CTGF mRNA 6 h after stimulation by VEGF. E (BREC) and F (BRPC), time-dependent induction of CTGF mRNA after stimulation by 25 ng/ml PlGF. Size of mRNAs corresponded to ~2.4 (CTGF) and ~1.6 kb (36B4). Representative Northern blots and control 36B4 (top) and quantitation of three experiments after normalization to the control signals are shown (bottom). Asterisks indicate significant differences at  $p < 0.05$  (\*),  $p < 0.01$  (\*\*), and  $p < 0.001$  (\*\*\*).

## RESULTS

**CTGF mRNA Expression by VEGF and PlGF**—The effects of VEGF on the expression of CTGF mRNA were studied by Northern blot analysis in BREC and BRPC. As shown in Fig. 1, A and B, 25 ng/ml VEGF increased CTGF mRNA (~2.4 kb) levels in a time-dependent manner, reaching a maximum after 6 h in BREC ( $3.1 \pm 0.70$ -fold,  $p < 0.001$ , Fig. 1A) and after 9 h in BRPC ( $2.0 \pm 0.22$ -fold,  $p < 0.01$ , Fig. 1B).

The dose response to VEGF-induced CTGF mRNA expression was studied after 6 h of VEGF stimulation. As shown in Fig. 1, C and D, the expression of CTGF mRNA was up-regulated in a dose-dependent manner, with significant increases observed at concentrations as low as 0.25 ng/ml in both BREC (Fig. 1C) and BRPC (Fig. 1D). Maximal increases were observed at VEGF concentrations of 25 ng/ml in both BREC and BRPC.

Since BREC and BRPC may express both KDR and Flt1, we examined the effects of PlGF, a Flt1-specific ligand (46–48), on the induction of CTGF gene expression in vascular cells. As shown in Fig. 1E, CTGF mRNA levels were not affected after stimulation of 25 ng/ml of PlGF in BREC. In contrast, PlGF increased CTGF mRNA after 3 h of stimulation, which peaked after 9 h in BRPC ( $1.9 \pm 0.30$ -fold,  $p < 0.01$ , Fig. 1F), suggest-

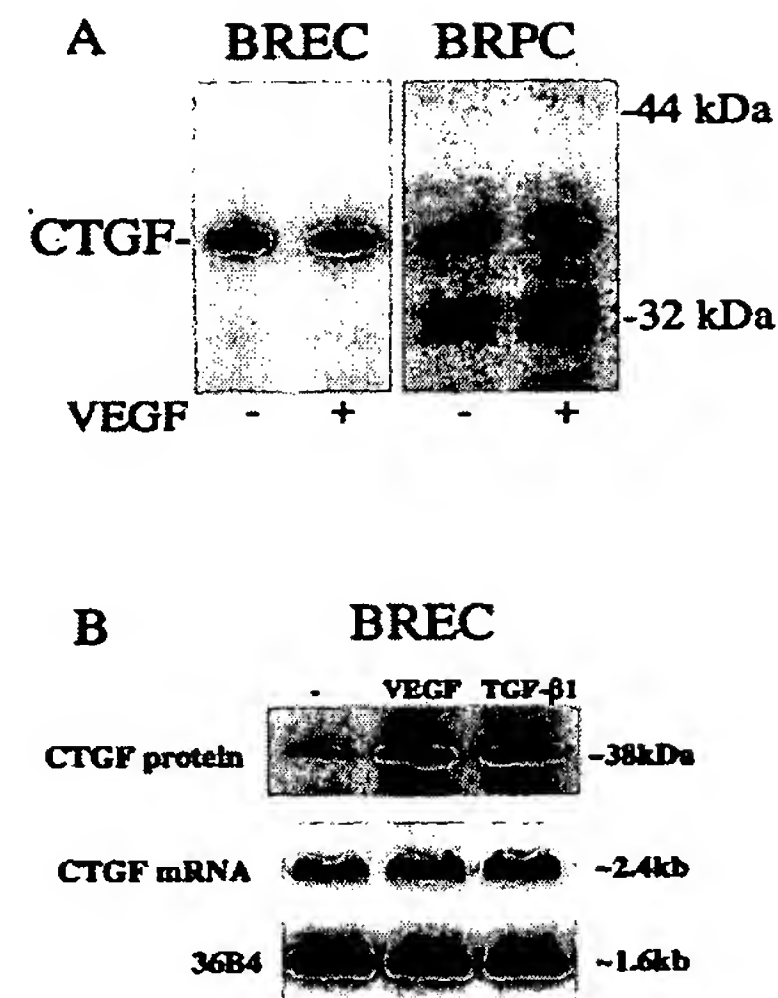


FIG. 2. Effect of VEGF on CTGF protein expression. A, confluent monolayers of BREC (left) or BRPC (right) in serum-free media were studied after 10 h, with and without 25 ng/ml of VEGF. An equal amount of protein for each sample was subjected to 10% SDS-gel electrophoresis and transferred to a nitrocellulose membrane. Signals were detected by Western blot analysis with anti-CTGF antibody. The size of the CTGF protein corresponded to ~38 kDa. B, VEGF (25 ng/ml) and TGF-β1 (10 ng/ml) were studied as described in Figs. 1 and 2A. VEGF and TGF-β1 increased CTGF protein and mRNA expression by a similar amount.

ing that VEGF-induced CTGF gene expression was mediated primarily by KDR in BREC and Flt1 in BRPC.

**VEGF Induction of CTGF Protein Production**—To determine if the effects of VEGF on CTGF mRNA were correlated with its protein level, CTGF protein expression was assessed by Western blot analysis using anti-human CTGF antibody. As shown in Fig. 2A, the detected size of CTGF protein was ~38 kDa in both BREC and BRPC. VEGF (25 ng/ml) increased the level of CTGF protein after 10 h in both BREC and BRPC. Comparative studies were performed on the effects of VEGF (25 ng/ml) and TGF-β1 (10 ng/ml) on the expression of CTGF mRNA and protein. As shown Fig. 2B, VEGF and TGF-β1 increased CTGF protein expression by a similar amount ( $2.5 \pm 0.4$ - and  $2.8 \pm 0.8$ -fold, respectively, in BREC). CTGF mRNA levels were also increased a similar extent ( $3.0 \pm 0.3$ - and  $3.3 \pm 0.5$ -fold, respectively).

**Effects of VEGF on the Half-life of CTGF mRNA**—The effects of VEGF on the stability of CTGF mRNA were examined. Northern blot analyses were performed with addition of actinomycin D (5 μg/ml) after 6 h of VEGF (25 ng/ml) stimulation. In BREC (Fig. 3A) and BRPC (Fig. 3B), the half-life of CTGF mRNA was 1.7 and 3.6 h, respectively. There was no significant difference between VEGF-treated and -untreated cells.

**Effects of Cycloheximide on CTGF mRNA Regulation**—In order to examine the possibility that VEGF regulates CTGF mRNA expression through new protein synthesis of cytokines or transcription factors, cells were treated for 6 h with VEGF (25 ng/ml) and a protein synthesis inhibitor, cycloheximide (10 μg/ml). Fig. 3, C and D, shows that cycloheximide did not prevent the increase of CTGF mRNA. Addition of both VEGF and cycloheximide increased CTGF mRNA  $2.4 \pm 0.41$ -fold in BREC (Fig. 3C) and  $2.5 \pm 0.40$ -fold in BRPC (Fig. 3D) after 6 h as compared with cycloheximide alone ( $p < 0.01$ ). These data suggest that the stimulation of CTGF mRNA expression by VEGF was not induced by increased synthesis of a regulatory protein.

**Involvement of ERK and PI3-Kinase-Akt in VEGF Signal-**



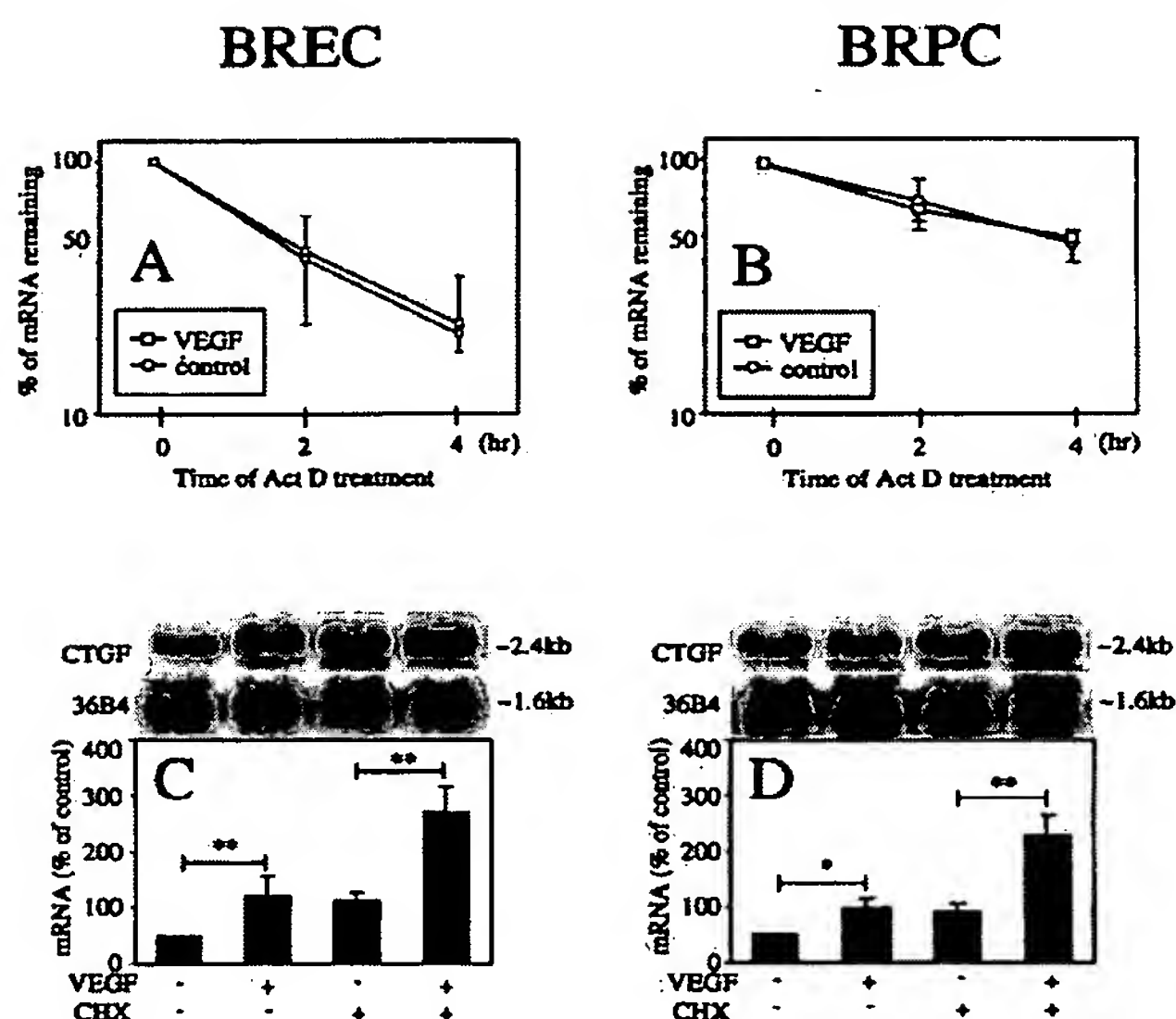


FIG. 3. *A* and *B*, decay of CTGF mRNA in the presence of actinomycin D. *C* and *D*, effects of cycloheximide on CTGF mRNA regulation. *A* (BREC) and *B* (BRPC), cells were incubated with or without 25 ng/ml VEGF for 6 h, and *de novo* mRNA transcription was inhibited by addition of actinomycin D (5  $\mu$ g/ml). Total RNA was extracted at 2 and 4 h after administration of actinomycin D. The values in the graph indicate the percentage of initial CTGF mRNA signal remaining in the specific conditions. *C* (BREC) and *D* (BRPC), cells were incubated with or without VEGF (25 ng/ml) and cycloheximide (CHX, 10  $\mu$ g/ml) for 6 h, and total RNA was analyzed. Representative Northern blots and control 36B4 (*top*) and quantitation of three experiments after normalization to the control signal are shown (*bottom*). Asterisks indicate significant differences at  $p < 0.05$  (\*) and  $p < 0.01$  (\*\*).

ing—Since ERK and PI3-kinase-Akt pathways have been reported to play central roles in VEGF signaling and biological actions (9–15), we investigated whether or not VEGF can activate ERK and PI3-kinase-Akt pathways equally in BREC and BRPC. As shown in Fig. 4A, immunoblot analysis of immunoprecipitates of KDR from BREC stimulated with VEGF or PlGF using an antibody to phosphotyrosine and PI3-kinase p85 subunit demonstrated that VEGF, but not PlGF, promoted the tyrosine phosphorylation of KDR and interactions of KDR and p85 subunit of PI3-kinase. In contrast, as shown in Fig. 4B, immunoblot analysis of immunoprecipitates of Flt1 from BRPC stimulated with VEGF or PlGF demonstrated that both VEGF and PlGF increased the tyrosine phosphorylation of Flt1 and interactions of Flt1 and p85 subunit of PI3-kinase. These data suggest that VEGF can activate the receptor tyrosine phosphorylation and interaction with PI3-kinase p85 subunit in both KDR and Flt1.

To investigate the activation of Akt and ERK, we next performed immunoblot analysis with anti-phosphorylated Akt or anti-phosphorylated ERK antibodies using total cell lysates from BREC or BRPC stimulated with VEGF. As shown in Fig. 4, *C* and *D*, VEGF induced phosphorylation of both Akt and ERK in BREC by 3.1- and 5.8-fold (Fig. 4C), but only induced phosphorylation of Akt in BRPC by 2.6-fold (Fig. 4D). No effect on ERK phosphorylation was observed in BRPC. These data suggest that VEGF activated both ERK and PI3-kinase-Akt pathways in BREC, but stimulated only PI3-kinase-Akt pathway in BRPC.

Since the activation of PI3-kinase by VEGF has not been reported in BRPC, we studied the effects of VEGF on PI3-kinase activity in BRPC. As shown in Fig. 5, the addition of VEGF (25 ng/ml) increased PI3-kinase activity in a time-dependent manner by  $2.1 \pm 0.27$ -fold ( $p < 0.01$ ) after 5 min and

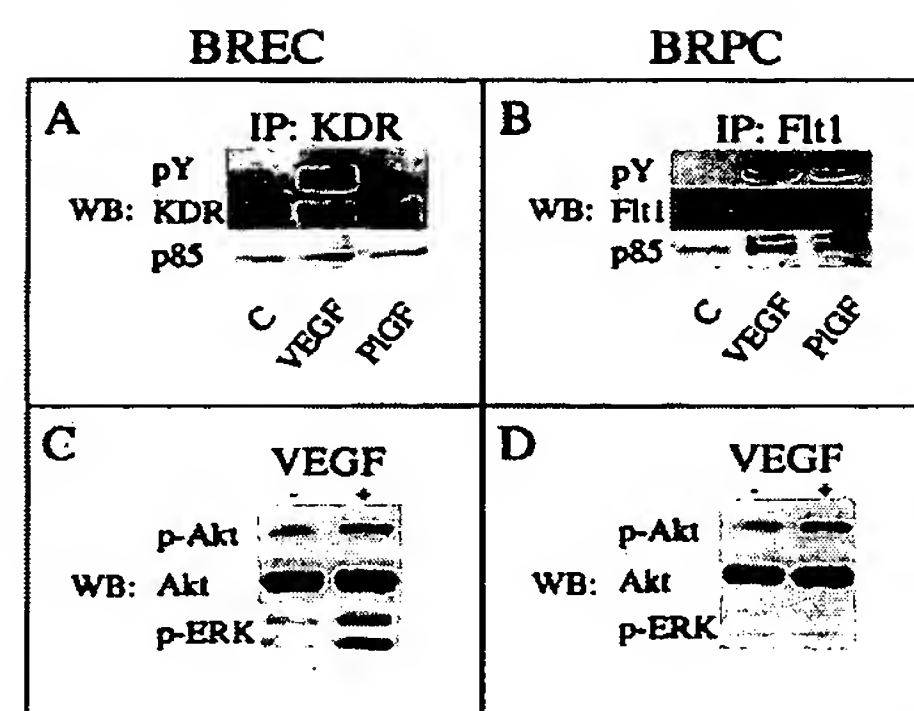


FIG. 4. VEGF effects on the ERK and PI3-kinase-Akt. *A* (BREC) and *B* (BRPC), cells were incubated with VEGF (25 ng/ml) or PlGF (25 ng/ml) for 5 min. Proteins immunoprecipitated (IP) with KDR (*A*) or Flt1 (*B*) antibodies were fractionated by SDS-polyacrylamide gel electrophoresis. Immunoblots were probed with an antibody to phosphotyrosine (pY, *top*) and reprobbed with antibody to p85 (*bottom*) and KDR (*A*, *middle*) or Flt1 (*B*, *middle*). *C* (BREC) and *D* (BRPC), cells were incubated with VEGF (25 ng/ml) for 15 min. Immunoblots were probed with an antibody to phosphorylated Akt (*top*) and reprobbed with antibody to total Akt (*middle*) or phosphorylated ERK (*bottom*). WB, Western blot.

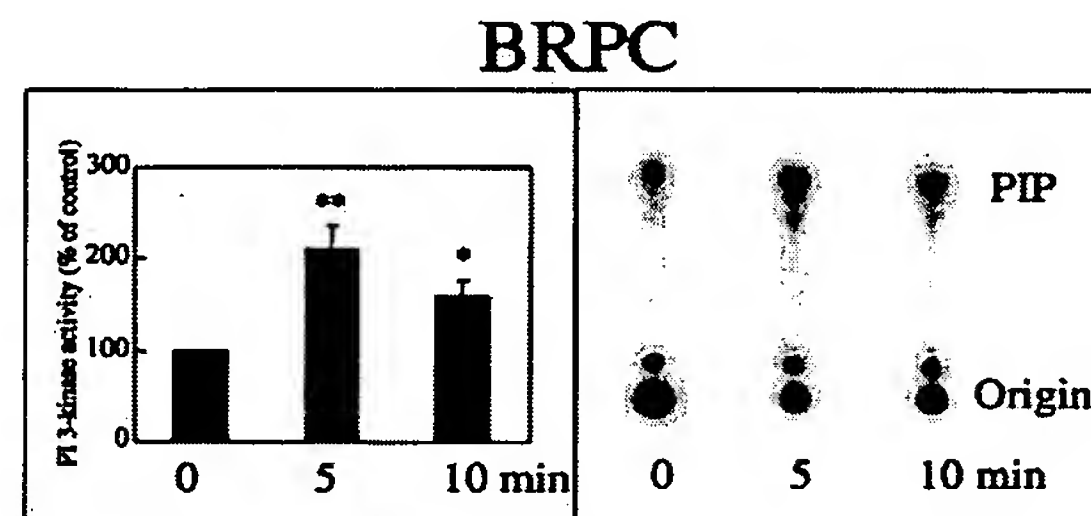


FIG. 5. Effects of VEGF on PI3-kinase activation in BRPC. BRPC were incubated with VEGF (25 ng/ml) for the indicated times and harvested. Equal amounts of lysates were immunoprecipitated with anti-phosphotyrosine antibody, and immunocomplexes were assayed for their ability to phosphorylate PI to phosphatidylinositol phosphate (PIP). Representative autoradiogram (*right*) and quantitation of three experiments in the percentage of intensity of control (*left*) are shown. Asterisks indicate significant differences at  $p < 0.05$  (\*) and  $p < 0.01$  (\*\*).

by  $1.6 \pm 0.17$ -fold ( $p < 0.05$ ) after 10 min in BRPC.

**Effects of PKC, ERK, and PI3-Kinase Inhibition on VEGF-induced CTGF Expression**—To investigate the signaling pathways involved in VEGF-induced CTGF expression, the effects of inhibition of PKC, ERK, and PI3-kinase were determined. Cells were treated with 25 ng/ml VEGF for 6 h after pretreatment with the kinase inhibitor GF 109203X, a classical and novel PKC-specific inhibitor (1  $\mu$ M) (49, 50); PD98059, a MAPK/ERK kinase inhibitor (20  $\mu$ M); or wortmannin, a PI3-kinase inhibitor (100 nM) (Fig. 6, *A* and *B*). Neither GF 109203X nor PD98059 had significant effects on VEGF-induced CTGF mRNA expression, but wortmannin inhibited the effects of VEGF by  $88 \pm 6.5\%$  ( $p < 0.01$ ) in BREC (Fig. 6A) and  $78 \pm 22\%$  ( $p < 0.01$ ) in BRPC (Fig. 6B). To confirm further the involvement of PI3-kinase in VEGF-induced CTGF expression, we used recombinant adenoviruses encoding DN Ras, DNERK, or  $\Delta$ p85 of PI3-kinase. BREC were transfected with each adenoviral vector, followed by stimulation with 25 ng/ml VEGF for 6 h. Neither DN Ras nor DNERK had significant effects on VEGF-induced increase in CTGF mRNA, but  $\Delta$ p85 of PI3-kinase completely inhibited VEGF-induced CTGF expression ( $p < 0.001$ , Fig. 6C).

**Role of PKC $\zeta$  and Akt/PKB in VEGF-induced CTGF Expression**—Since it has been reported that atypical PKC (51–55) and



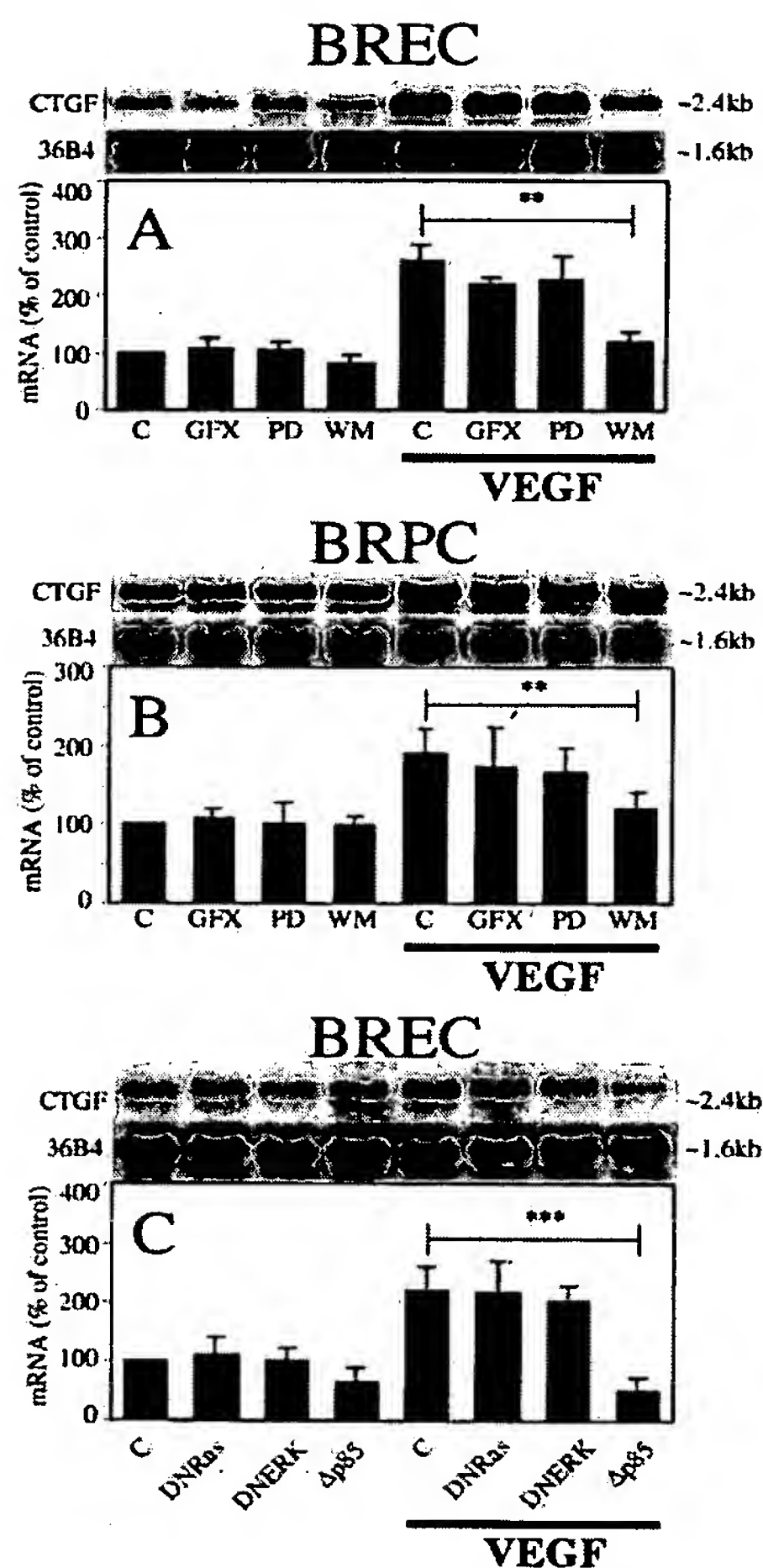


FIG. 6. Characterization of the role of PKC, ERK, and PI3-kinase in VEGF-induced CTGF expression. A (BREC) and B (BRPC), cells were pretreated with GF 109203X (GFX, 1  $\mu$ M), PD98059 (PD, 20  $\mu$ M), or wortmannin (WM, 100 nM) followed by stimulation with 25 ng/ml VEGF for 6 h. C, BREC were infected with control (C), DNRas, DNERK, or  $\Delta$ p85 by adenoviral vectors. After 24 h, cells were starved overnight, followed by stimulation with 25 ng/ml VEGF for 6 h. Representative Northern blots and control 36B4 (top) and quantitation of three experiments after normalization to the control signals are shown (bottom). Asterisks indicate significant differences at  $p < 0.01$  (\*\*) and  $p < 0.001$  (\*\*\*).

Akt/PKB (13, 14, 56) have significant roles as signaling molecules downstream of PI3-kinase, we examined the involvement of PKC $\zeta$  and Akt in this process. BREC were infected with each adenoviral vector, followed by stimulation with 25 ng/ml VEGF for 6 h. Neither wild type PKC $\zeta$  nor DNPCK $\zeta$  had significant effects on VEGF-induced increase in CTGF mRNA (Fig. 7A). In contrast, as shown in Fig. 7B, infection with CAAkt increased CTGF mRNA expression  $2.1 \pm 0.21$ -fold ( $p < 0.01$ ) without VEGF and  $2.5 \pm 0.40$ -fold with VEGF. Overexpression with adenoviral vector containing DNAkt inhibited VEGF-induced CTGF expression by  $85 \pm 13\%$  ( $p < 0.01$ ).

#### DISCUSSION

In this study, we have shown that VEGF can increase the mRNA expression of CTGF in a time- and concentration-dependent manner in both microvascular endothelial cells and contractile cells (capillary pericytes) possibly indicating that the effects of VEGF on CTGF expression may occur in all cells with VEGF receptors. This possibility is supported by the re-

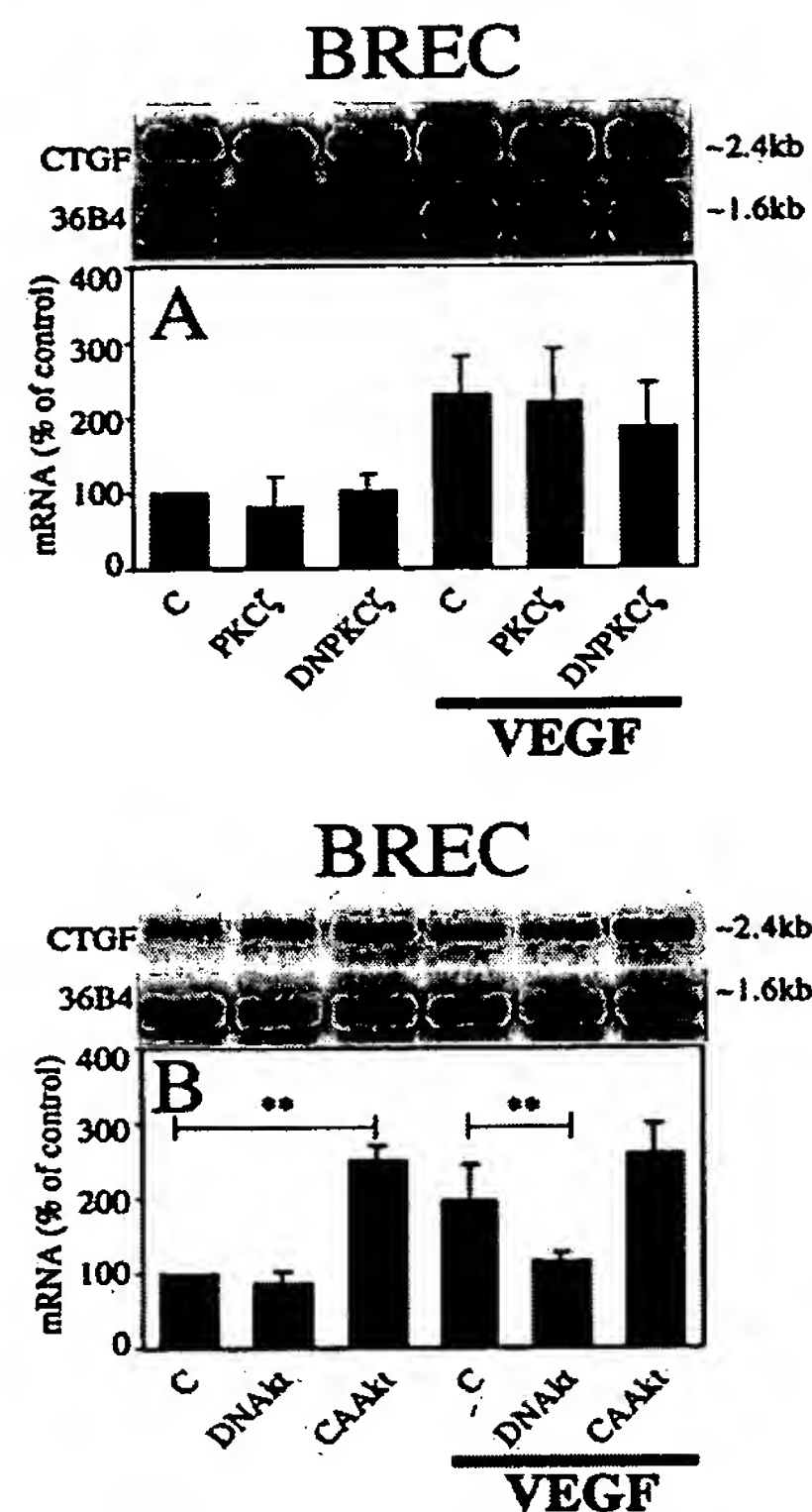


FIG. 7. Role of PKC $\zeta$  and Akt/PKB in VEGF-induced CTGF expression. BREC were infected with adenoviral vectors containing cDNAs of control, wild type PKC $\zeta$ , and DNPCK $\zeta$  (A) or control, DNAkt, and CAAkt (B). After 24 h, cells were starved overnight, followed by stimulation with 25 ng/ml VEGF for 6 h. Representative Northern blots and control 36B4 (top) and quantitation of three experiments after normalization to the control signals are shown (bottom). Asterisks indicate significant differences at  $p < 0.01$  (\*\*).

sults showing that both Flt1 (VEGFR1) and KDR/Flk1 (VEGFR2) can mediate the increases in CTGF mRNA expression. The ability of Flt1 to induce increases in CTGF mRNA levels is demonstrated in the pericytes that have predominantly Flt1 receptors and where the expression of KDR/Flk1 receptors were not significantly high enough to be determined by Northern blot analysis as reported in a previous publication (57). In addition, PlGF, a Flt1 receptor-specific ligand (48), was able to induce CTGF mRNA levels in BRPC but not in BREC, again supporting the postulate that VEGF can induce CTGF mRNA by activating through Flt1 in pericytes. The KDR/Flk1 receptors in the endothelial cells can also induce CTGF gene expression since KDR/Flk1 receptors are the predominant VEGF receptors in endothelial cells (58), and PlGF was not effective in inducing CTGF mRNA expression in endothelial cells. Further studies will be necessary to determine whether other types of VEGF receptors, such as Flt4 (59) and neuropilin-1 (60), which are present in endothelial cells, can also induce CTGF expression.

The VEGF dose-response curves for CTGF in both BRPC and BREC are similar and suggest that VEGF binds to high affinity receptors, consistent with the known  $K_d$  values of Flt1 and KDR/Flk1 at 10–100 pM (58, 61). VEGF-induced CTGF mRNA is most likely due to an induction of transcription rather than altering the half-life of CTGF mRNA since the addition of VEGF failed to change the degradation rates of CTGF mRNA. The time course of the action of VEGF on CTGF (which required 6–9 h) suggests this is potentially a chronic action of VEGF. In addition, the time needed to achieve maximum effect

is also consistent with the calculated mRNA half-life of CTGF mRNA of 2–4 h.

From a biological perspective, the effects of VEGF on CTGF mRNA could potentially have important physiological impact for several reasons. First is that the increase in CTGF mRNA results in increased protein levels. Second, the VEGF concentration that was minimally active (0.25 ng/ml) can easily bind and activate a significant percentage of the VEGFR-1, -2 receptors (58, 61). Third, this low level of VEGF may exist even in non-pathological states, suggesting that low levels of VEGF may have physiological actions on maintaining extracellular matrix production via the induction of CTGF. At 2.5–25 ng/ml VEGF which are encountered in hypoxic and angiogenic states (62), the induction of CTGF expression by VEGF could potentially induce the fibrosis that frequently accompanies neovascularization. This possibility is supported further by the demonstration that the protein levels of CTGF expression were increased 10 h after the addition of VEGF that was consistent with the maximum increase in the mRNA levels at 6–9 h. In addition, the potency of VEGF on CTGF expression appeared to be similar to TGF- $\beta$ 1, suggesting that both of them could induce fibrosis associated with neovascularization.

The activation of the endogenous tyrosine kinases of KDR/Flk1s can stimulate multiple signaling pathways, including Ras-ERK (63), PI3-kinase-Akt (13–15), and phospholipase C- $\gamma$ -PKC (9–11) cascades. Much less is known regarding the regulation of Flt1 receptors. The results in BREC confirmed previous publications that VEGF can increase the tyrosine phosphorylation of KDR/Flk1 and its interaction with p85 subunit of PI3-kinase. In addition, VEGF also activates the ERK1/2 pathway confirming earlier reports from many laboratories, including ours. In contrast, VEGF was unable to activate ERK1/2 but stimulated the activation of PI3-kinase and phosphorylation of Akt in BRPC. These results have provided further direct evidence that the signaling pathways for Flt1 in vascular cells are different from those for KDR/Flk1. The lack of effect on ERK1/2 activation also supports the hypothesis that Flt1, unlike KDR/Flk1, is not involved in mitogenic actions (64). Further studies will be needed to determine the structural differences responsible for the inability of VEGFR1 to engage the Ras-ERK pathway.

The present report does provide the strong evidence that VEGF is inducing CTGF gene expression in both endothelial cells and pericytes via VEGFR1 or -R2 by the activation of PI3-kinase and Akt. This evidence includes the ability of wortmannin, a PI3-kinase inhibitor, to inhibit the effects of VEGFs in both cell types, whereas PD98059, a MAPK/ERK kinase inhibitor, and GF109203X, a general classical PKC and novel PKC inhibitor (1  $\mu$ M) (49, 50), did not have significant actions. Adenovirus containing dominant negative mutants of p85 subunit of PI3-kinase or Akt inhibited the action of VEGFs, whereas overexpression of dominant negative mutants of Ras and ERK1 by adenovirus vectors did not inhibit CTGF mRNA expression. Conversely, the overexpression of constitutive active Akt increased CTGF mRNA expression by 2.5-fold. The molecular steps between Akt activation and the enhancement of CTGF gene expression in the nucleus remain unclear, although the PKC $\zeta$  isoform is most likely not involved since the overexpression of either the wild type or dominant negative of PKC $\zeta$  isoform did not alter the effects of VEGF on CTGF mRNA levels.

The molecular processes between Akt phosphorylation and CTGF gene expression in the nucleus have not been studied. However, Pendurthi *et al.* (31) reported that factor VII and thrombin induced CTGF gene expression through a PI3-kinase-dependent pathway. TGF- $\beta$  has been reported to increase the

transcription rates of CTGF. A promoter element of CTGF, which is responsive to TGF- $\beta$  stimulation, has been reported to be present between –162 and –128 nucleotides in the 5' region (65). However, it is unlikely that the effects of VEGF on CTGF mRNA levels are mediated via the expression of TGF- $\beta$  since the addition of cycloheximide did not change these effects.

In summary, these results have provided the first evidence that VEGF can induce the expression of CTGF via both Flt1 and KDR/Flk1 by the selectively activated PI3-kinase-Akt pathway but independent of the Ras-ERK pathway. In addition, the spectrum of signaling pathways may be different between Flt1 and KDR/Flk1, possibly reflecting their physiological roles. Biologically, these results support the conclusion that VEGF, through its effects on CTGF expression, may have physiological roles such as the maintenance of capillary strength and wound healing via the extracellular matrix production. In disease states, VEGF-induced CTGF may cause the proliferation of fibrocellular components in retinal neovascular diseases such as proliferative diabetic retinopathy and age-related macular degeneration.

**Acknowledgment**—We thank Dr. Edward P. Feener for suggestions during the preparation of this manuscript.

#### REFERENCES

- Klagsbrun, M., and D'Amore, P. A. (1991) *Annu. Rev. Physiol.* **53**, 217–239
- Aiello, L. P., Gardner, T. W., King, G. L., Blankenship, G., Cavallerano, J. D., Ferris, F. L., III, and Klein, R. (1998) *Diabetes Care* **21**, 143–156
- Firestein, G. S. (1999) *J. Clin. Invest.* **103**, 3–4
- Lopez, P. F., Sippy, B. D., Lambert, H. M., Thach, A. B., and Hinton, D. R. (1996) *Invest. Ophthalmol. & Visual Sci.* **37**, 855–868
- Neufeld, G., Cohen, T., Gengrinovitch, S., and Poltorak, Z. (1999) *FASEB J.* **13**, 9–22
- Shweiki, D., Itin, A., Soffer, D., and Keshet, E. (1992) *Nature* **359**, 843–845
- Ferrara, N., Carver-Moore, K., Chen, H., Dowd, M., Lu, L., O'Shea, K. S., Powell-Braxton, L., Hillan, K. J., and Moore, M. W. (1996) *Nature* **380**, 439–442
- Petrova, T. V., Makinen, T., and Alitalo, K. (1999) *Exp. Cell Res.* **253**, 117–130
- Wu, L. W., Mayo, L. D., Dunbar, J. D., Kessler, K. M., Baerwald, M. R., Jaffe, E. A., Wang, D., Warren, R. S., and Donner, D. B. (2000) *J. Biol. Chem.* **275**, 5096–5103
- Xia, P., Aiello, L. P., Ishii, H., Jiang, Z. Y., Park, D. J., Robinson, G. S., Takagi, H., Newsome, W. P., Jirousek, M. R., and King, G. L. (1996) *J. Clin. Invest.* **98**, 2018–2026
- Takahashi, T., Ueno, H., and Shibuya, M. (1999) *Oncogene* **18**, 2221–2230
- He, H., Venema, V. J., Gu, X., Venema, R. C., Marrero, M. B., and Caldwell, R. B. (1999) *J. Biol. Chem.* **274**, 25130–25135
- Fujio, Y., and Walsh, K. (1999) *J. Biol. Chem.* **274**, 16349–16354
- Gerber, H. P., McMurtrey, A., Kowalski, J., Yan, M., Keyt, B. A., Dixit, V., and Ferrara, N. (1998) *J. Biol. Chem.* **273**, 30336–30343
- Thakker, G. D., Hajjar, D. P., Muller, W. A., and Rosengart, T. K. (1999) *J. Biol. Chem.* **274**, 10002–10007
- Bork, P. (1993) *FEBS Lett.* **327**, 125–130
- Lau, L. F., and Lam, S. C. (1999) *Exp. Cell Res.* **248**, 44–57
- Frazier, K., Williams, S., Kothapalli, D., Klapper, H., and Grotendorst, G. R. (1996) *J. Invest. Dermatol.* **107**, 404–411
- Kireeva, M. L., Latinkic, B. V., Kolesnikova, T. V., Chen, C. C., Yang, G. P., Abler, A. S., and Lau, L. F. (1997) *Exp. Cell Res.* **233**, 63–77
- Bradham, D. M., Igarashi, A., Potter, R. L., and Grotendorst, G. R. (1991) *J. Cell Biol.* **114**, 1285–1294
- Igarashi, A., Okochi, H., Bradham, D. M., and Grotendorst, G. R. (1993) *Mol. Biol. Cell* **4**, 637–645
- Kothapalli, D., Frazier, K. S., Welply, A., Segarini, P. R., and Grotendorst, G. R. (1997) *Cell Growth Differ.* **8**, 61–68
- Igarashi, A., Nashiro, K., Kikuchi, K., Sato, S., Ihn, H., Fujimoto, M., Grotendorst, G. R., and Takehara, K. (1996) *J. Invest. Dermatol.* **106**, 729–733
- Frazier, K. S., and Grotendorst, G. R. (1997) *Int. J. Biochem. Cell Biol.* **29**, 153–161
- Oemar, B. S., Werner, A., Garnier, J. M., Do, D. D., Godoy, N., Nauck, M., Marz, W., Rupp, J., Pech, M., and Luscher, T. F. (1997) *Circulation* **95**, 831–839
- Babic, A. M., Chen, C. C., and Lau, L. F. (1999) *Mol. Cell. Biol.* **19**, 2958–2966
- Shimo, T., Nakanishi, T., Nishida, T., Asano, M., Kanyama, M., Kuboki, T., Tamatani, T., Tezuka, K., Takemura, M., Matsumura, T., and Takigawa, M. (1999) *J. Biochem. (Tokyo)* **126**, 137–145
- Dammeier, J., Beer, H. D., Brauchle, M., and Werner, S. (1998) *J. Biol. Chem.* **273**, 18185–18190
- Murphy, M., Godson, C., Cannon, S., Kato, S., Mackenzie, H. S., Martin, F., and Brady, H. R. (1999) *J. Biol. Chem.* **274**, 5830–5834
- Ricupero, D. A., Romero, J. R., Rishikof, D. C., and Goldstein, R. H. (2000) *J. Biol. Chem.* **275**, 12475–12480
- Pendurthi, U. R., Allen, K. E., Ezban, M., and Rao, L. V. (2000) *J. Biol. Chem.* **275**, 14632–14641
- Abraham, D. J., Shiwen, X., Black, C. M., Sa, S., Xu, Y., and Leask, A. (2000)

- J. Biol. Chem.* 275, 15220–15225
33. Boes, M., Dake, B. L., Booth, B. A., Erondy, N. E., Oh, Y., Hwa, V., Rosenfeld, R., and Bar, R. S. (1999) *Endocrinology* 140, 1575–1580
  34. King, G. L., Goodman, A. D., Buzney, S., Moses, A., and Kahn, C. R. (1985) *J. Clin. Invest.* 75, 1028–1036
  35. Nayak, R. C., Berman, A. B., George, K. L., Eisenbarth, G. S., and King, G. L. (1988) *J. Exp. Med.* 167, 1003–1015
  36. Burgering, B. M., and Coffey, P. J. (1995) *Nature* 376, 599–602
  37. Kitamura, T., Ogawa, W., Sakaue, H., Hino, Y., Kuroda, S., Takata, M., Matsumoto, M., Maeda, T., Konishi, H., Kikkawa, U., and Kasuga, M. (1998) *Mol. Cell. Biol.* 18, 3708–3717
  38. Ueki, K., Yamamoto-Honda, R., Kaburagi, Y., Yamauchi, T., Tobe, K., Burgering, B. M., Coffey, P. J., Komuro, I., Akanuma, Y., Yazaki, Y., and Kadowaki, T. (1998) *J. Biol. Chem.* 273, 5315–5322
  39. Her, J. H., Lakhani, S., Zu, K., Vila, J., Dent, P., Sturgill, T. W., and Weber, M. J. (1993) *Biochem. J.* 296, 25–31
  40. Hara, K., Yonezawa, K., Sakaue, H., Ando, A., Kotani, K., Kitamura, T., Kitamura, Y., Ueda, H., Stephens, L., Jackson, T. R., Hawkins, P. T., Dhand, R., Clark, A. E., Holman, G. D., Waterfield, M. D., and Kasuga, M. (1994) *Proc. Natl. Acad. Sci. U. S. A.* 91, 7415–7419
  41. Ueberall, F., Hellbert, K., Kampfer, S., Maly, K., Villunger, A., Spitaler, M., Mwanjewe, J., Baier-Bitterlich, G., Baier, G., and Grunicke, H. H. (1999) *J. Cell Biol.* 144, 413–425
  42. He, T. C., Zhou, S., da Costa, L. T., Yu, J., Kinzler, K. W., and Vogelstein, B. (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 2509–2514
  43. Miyake, S., Makimura, M., Kanegae, Y., Harada, S., Sato, Y., Takamori, K., Tokuda, C., and Saito, I. (1996) *Proc. Natl. Acad. Sci. U. S. A.* 93, 1320–1324
  44. Kuboki, K., Jiang, Z. Y., Takahara, N., Ha, S. W., Igarashi, M., Yamauchi, T., Feener, E. P., Herbert, T. P., Rhodes, C. J., and King, G. L. (2000) *Circulation* 101, 676–681
  45. Liang, P., and Pardee, A. B. (1992) *Science* 257, 967–971
  46. Clauss, M., Weich, H., Breier, G., Knies, U., Rockl, W., Waltenberger, J., and Risau, W. (1996) *J. Biol. Chem.* 271, 17629–17634
  47. Sawano, A., Takahashi, T., Yamaguchi, S., Aonuma, M., and Shibuya, M. (1996) *Cell Growth Differ.* 7, 213–221
  48. Park, J. E., Chen, H. H., Winer, J., Houck, K. A., and Ferrara, N. (1994) *J. Biol. Chem.* 269, 25646–2554
  49. Park, J. Y., Takahara, N., Gabriele, A., Chou, E., Naruse, K., Suzuma, K., Yamauchi, T., Ha, S. W., Meier, M., Rhodes, C. J., and King, G. L. (2000) *Diabetes* 49, 1239–1248
  50. Toullec, D., Pianetti, P., Coste, H., Bellevergue, P., Grand-Perret, T., Ajakane, M., Baudet, V., Boissin, P., Boursier, E., Loriolle, F., Dunhanel, L., Charon, D., and Kirilovsky, J. (1991) *J. Biol. Chem.* 266, 15771–15781
  51. Kotani, K., Ogawa, W., Matsumoto, M., Kitamura, T., Sakaue, H., Hino, Y., Miyake, K., Sano, W., Akimoto, K., Ohno, S., and Kasuga, M. (1998) *Mol. Cell. Biol.* 18, 6971–6982
  52. Le Good, J. A., Ziegler, W. H., Parekh, D. B., Alessi, D. R., Cohen, P., and Parker, P. J. (1998) *Science* 281, 2042–2045
  53. Standaert, M. L., Galloway, L., Karnam, P., Bandyopadhyay, G., Moscat, J., and Farese, R. V. (1997) *J. Biol. Chem.* 272, 30075–30082
  54. Standaert, M. L., Bandyopadhyay, G., Perez, L., Price, D., Galloway, L., Poklepovic, A., Sajan, M. P., Cenni, V., Sirri, A., Moscat, J., Toker, A., and Farese, R. V. (1999) *J. Biol. Chem.* 274, 25308–25316
  55. Wellner, M., Maasch, C., Kupprion, C., Lindschau, C., Luft, F. C., and Haller, H. (1999) *Arterioscler. Thromb. Vasc. Biol.* 19, 178–185
  56. Jiang, B. H., Zheng, J. Z., Aoki, M., and Vogt, P. K. (2000) *Proc. Natl. Acad. Sci. U. S. A.* 97, 1749–1753
  57. Takagi, H., King, G. L., and Aiello, L. P. (1996) *Diabetes* 45, 1016–1023
  58. Millauer, B., Witzmann-Voos, S., Schnurch, H., Martinez, R., Moller, N. P., Risau, W., and Ullrich, A. (1993) *Cell* 72, 835–846
  59. Joukov, V., Pajusola, K., Kaipainen, A., Chilov, D., Lahtinen, I., Kukk, E., Saksela, O., Kalkkinen, N., and Alitalo, K. (1996) *EMBO J.* 15, 290–298
  60. Soker, S., Takashima, S., Miao, H. Q., Neufeld, G., and Klagsbrun, M. (1998) *Cell* 92, 735–745
  61. de-Vries, C., Escobedo, J. A., Ueno, H., Houck, K., Ferrara, N., and Williams, L. T. (1992) *Science* 255, 989–991
  62. Aiello, L. P., Avery, R. L., Arrigg, P. G., Keyt, B. A., Jampel, H. D., Shah, S. T., Pasquale, L. R., Thieme, H., Iwamoto, M. A., Park, J. E., Nguyen, H. V., Aiello, L. M., Ferrara, N., and King, G. L. (1994) *N. Engl. J. Med.* 331, 1480–1487
  63. Guo, D., Jia, Q., Song, H. Y., Warren, R. S., and Donner, D. B. (1995) *J. Biol. Chem.* 270, 6729–6733
  64. Kanno, S., Oda, N., Abe, M., Terai, Y., Ito, M., Shitara, K., Tabayashi, K., Shibuya, M., and Sato, Y. (2000) *Oncogene* 19, 2138–2146
  65. Grotendorst, G. R., Okochi, H., and Hayashi, N. (1996) *Cell Growth Differ.* 7, 469–480



## Basic fibroblast growth factor is a substrate for protein phosphorylation and is phosphorylated by capillary endothelial cells in culture

(growth factors/cAMP-dependent protein kinase/protein kinase C)

JEAN-JACQUES FEIGE AND ANDREW BAIRD

Laboratories for Neuroendocrinology, The Salk Institute, La Jolla, CA 92037

Communicated by Roger Guillemin, December 12, 1988

**ABSTRACT** A phosphorylated basic fibroblast growth factor (FGF) can be detected in extracts of bovine capillary endothelial cells and human hepatoma cells. Accordingly, human basic FGF contains consensus sequences that account for its phosphorylation on Thr-112 by the catalytic subunit of the cAMP-dependent protein kinase A (PK-A) and on Ser-64 by the calcium- and phospholipid-dependent protein kinase C (PK-C). A kinetic analysis of both of these reactions revealed that basic FGF is among the better substrates for these enzymes. Although the kinase responsible for the phosphorylation *in vivo* has not yet been identified, we examined the effects of phosphorylation on the biological activity, heparin-binding capacity, and receptor-binding capacity of phosphorylated basic FGF. No effects of phosphorylation were observed when the mitogen was phosphorylated by PK-C. In contrast, when basic FGF was phosphorylated in the receptor-binding domain with PK-A, the growth factor was 3-8 times better at displacing radiolabeled basic FGF in the radioreceptor assay. No effects were seen on the binding of this FGF to immobilized heparin or cell-associated glycosaminoglycans, suggesting that this phosphorylation modifies the affinity of basic FGF for its receptor. Biological assays for basic FGF failed to identify differences between the phosphorylated and unphosphorylated forms of recombinant basic FGFs presumably because of the presence of ectophosphatases and the experimental conditions of proliferation and mitogenic assays (37°C, 24-96 hr). Because the relative affinity of basic FGF for its receptor and cell-associated glycosaminoglycans may regulate its activity, the identification of a modified form of basic FGF may be of particular importance in understanding the mechanisms that regulate its biological activity, bioavailability, and processing to and from the extracellular matrix.

Although the acidic and basic fibroblast growth factors (FGFs) have been clearly established to participate in the regulation of the proliferation of many cell types, the structural characterization of these molecules (1, 2) has raised many questions regarding their possible mechanism of action. As an example, both mitogens lack a classical signal sequence (3, 4) that would allow their secretion and access to their putative receptor located on the extracellular surface of the plasma membrane (5, 6). Because both acidic and basic FGFs have been found in the extracellular matrix, it has been suggested that they utilize an alternative secretion pathway to the outside of the cell and that their activity is tightly regulated by limiting their bioavailability (7-10). In view of the fact that both FGFs are also characterized by their high affinity for immobilized heparin and that they remain cell-associated even when expressed with consensus signal peptide sequences (11, 12), we have been investigating the

possibilities that these proteins are cytoplasmic and that posttranslational changes might modulate their processing to and from the extracellular matrix.

Protein phosphorylation is a posttranslational modification that has been implicated in the regulation of almost all steps of cell division (13-16). Recently, Plouët *et al.* (17) reported that the release of acidic FGF from rod outer segment membranes requires ATP-dependent protein phosphorylation. Because basic and acidic FGFs contain consensus sequences for phosphorylation by the phospholipid- and calcium-dependent protein kinase C (PK-C) and basic FGF contains a consensus sequence for the cAMP-dependent protein kinase A (PK-A), we tested the hypothesis that growth factors, and in particular FGFs, are potential substrates for protein phosphorylation.

### MATERIALS AND METHODS

**Materials.** PK-C was a gift from G. Walton (University of California-San Diego, La Jolla) and I. Vilgrain and J. M. Pelarin (Unité 244, Institut National de la Santé et de la Recherche Médicale, Grenoble, France). The catalytic subunit of PK-A was provided by S. Taylor (University of California-San Diego, La Jolla). Casein kinases I and II were purified from bovine lung as previously described and were provided to us by C. Cochet (Unité 244, Institut National de la Santé et de la Recherche Médicale, Grenoble) (18, 19). Recombinant human basic and acidic FGFs (20) were obtained from P. Barr and L. Cousens (Chiron). Synthetic transforming growth factor  $\alpha$  (TGF- $\alpha$ ) was supplied by Nicholas Ling (Salk Institute), and platelet-derived growth factor (PDGF) was a gift of R. Ross (University of Washington, Seattle). Recombinant insulin-like growth factor I (IGF-I) was obtained from Fujisawa Pharmaceuticals (Osawa, Japan); nerve growth factor (NGF), from D. Schubert (Salk Institute); and tumor necrosis factor  $\alpha$  (TNF- $\alpha$ ), from Suntory Institute for Biomedical Research (Osaka, Japan). [ $\gamma$ - $^{32}$ P]ATP (4000 Ci/mmol; 1 Ci = 37 GBq) and [ $\gamma$ - $^{32}$ P]orthophosphate were purchased from ICN. Antiserum 773 is an antiserum against basic FGF-(1-24) conjugated to bovine serum albumin and was raised in rabbits. Protein A-Sepharose was obtained from Pharmacia. Reagents for sodium dodecyl sulfate/polyacrylamide gel electrophoresis (SDS/PAGE) were obtained from Bio-Rad, and all other materials were from Sigma.

**Phosphorylation of Growth Factors by PK-C and PK-A.** Different growth factors ( $\approx 0.5 \mu\text{g}$ ) were incubated for 15 min at 30°C in the presence of purified PK-C from bovine brain

**Abbreviations:** FGF, fibroblast growth factor; PK-A, cAMP-dependent protein kinase; PK-C, calcium- and phospholipid-dependent protein kinase; IGF, insulin-like growth factor; TGF, transforming growth factor; PDGF, platelet-derived growth factor; EGF, epidermal growth factor; NGF, nerve growth factor; TNF, tumor necrosis factor; ACE cells, adrenocortical capillary endothelial cells.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

(21), the catalytic subunit of PK-A purified from porcine skeletal muscle (22), or bovine lung casein kinases I and II (18, 19). PK-C assays were performed in 20  $\mu$ l of 10 mM Tris-HCl (pH 7.5) containing 10  $\mu$ M [ $\gamma$ - $^{32}$ P]ATP (1500 cpm/pmol), 10 mM MgCl<sub>2</sub>, 0.6 mM CaCl<sub>2</sub>, 40  $\mu$ g of phosphatidylserine per ml, 0.8  $\mu$ g of dioctanoylglycerol per ml; and 1  $\mu$ g of PK-C per ml. PK-A assays were performed in 20  $\mu$ l of 10 mM Tris-HCl (pH 7.5) containing 5  $\mu$ M [ $\gamma$ - $^{32}$ P]ATP (3000 cpm/pmol), 10 mM MgCl<sub>2</sub>, and 3  $\mu$ g of the catalytic subunit of PK-A per ml. Casein kinase assays were performed in 20  $\mu$ l of 10 mM Tris-HCl (pH 7.5) containing 10  $\mu$ M [ $\gamma$ - $^{32}$ P]ATP (3000 cpm/pmol), 50 mM MgCl<sub>2</sub>, and either 10  $\mu$ g of CK-I or 4  $\mu$ g of CK-II per ml. The reactions were stopped with Laemmli sample buffer, and the phosphorylated proteins were separated by 15% or 20% SDS/PAGE (22) on a mini-slab-gel apparatus (Idea Scientific, Corvallis, OR) and visualized by overnight autoradiography.

**Phosphoamino Acid Analysis and Tryptic Mapping.** Radiolabeled peptides were extracted from the polyacrylamide gel in 0.05 M ammonium bicarbonate (pH 7.3–7.6) supplemented with 0.1% SDS and 1% 2-mercaptoethanol. After precipitation with 50% trichloroacetic acid, the pellet was dissolved in 6 M HCl, and the protein was hydrolyzed for 60 min at 110°C. Phosphoamino acids were separated by two-dimensional high-voltage electrophoresis as described (23). The plates were run for 20 min at 1.5 kV in pH 1.9 buffer in the first dimension and then for 16 min at 1.3 kV in pH 3.5 buffer in the second dimension.

**Phosphorylation of Basic FGF by Cells in Culture.** Three 10-cm dishes of SK-HEP-1 hepatoma cells and adrenocortical capillary endothelial cells (ACE cells) ( $21 \times 10^6$  cells) were labeled overnight at 37°C in phosphate-free Dulbecco's modified Eagle's medium (DMEM) supplemented with 2% dialyzed calf serum, 0.5% (SK-HEP-1) or 5% (ACE) normal DMEM, and 0.25  $\mu$ Ci (SK-HEP-1) or 0.40  $\mu$ Ci (ACE) of [ $^{32}$ P]orthophosphate per ml. At the end of the labeling period, cells were washed twice with phosphate-buffered saline and lysed in radioimmunoprecipitation assay (RIPA) buffer (24) for 30 min at 4°C. After centrifugation in an Eppendorf Microfuge, the supernatants were immunoprecipitated by using the polyclonal antibody 773 (1:200) in the presence or absence of 5  $\mu$ g of basic FGF. Twenty-five microliters of protein A-Sepharose suspension (50% in RIPA buffer) was added to the tube and mixed for 30 min. At the end of this incubation, the solution was centrifuged, and the pellet was washed three times with RIPA buffer, resuspended in Laemmli sample buffer and analyzed by 15% SDS/PAGE (22).

**Biological Assays of FGF Activity.** Basic FGF was phosphorylated with PK-C or the catalytic subunit of PK-A in the presence of unlabeled ATP under the conditions described above. Controls consisted of basic FGF treated with an identical incubation, but in the absence of enzyme. The phosphorylated and nonphosphorylated mitogens were tested for their capacity to bind the FGF receptor by using the BHK cell assay described by Moscatelli (25). Their effects on the proliferation of capillary endothelial cells was performed as described (26), and cell numbers were determined with a Coulter Counter.

## RESULTS

**Acidic and Basic FGFs Are Substrates for Protein Phosphorylation.** Human basic FGF was phosphorylated by both PK-C and the catalytic subunit of PK-A (Fig. 1). In the absence of substrate, there was clear self-phosphorylation of PK-C, PK-A, and casein kinase II (lanes 1, 2, and 4), and as expected (19, 21) casein kinase I showed little, if any, self-phosphorylation (lane 3). Basic FGF was a good substrate for PK-C (lane 5) as shown by the appearance of an intense radiolabeled band corresponding to 18 kDa. This

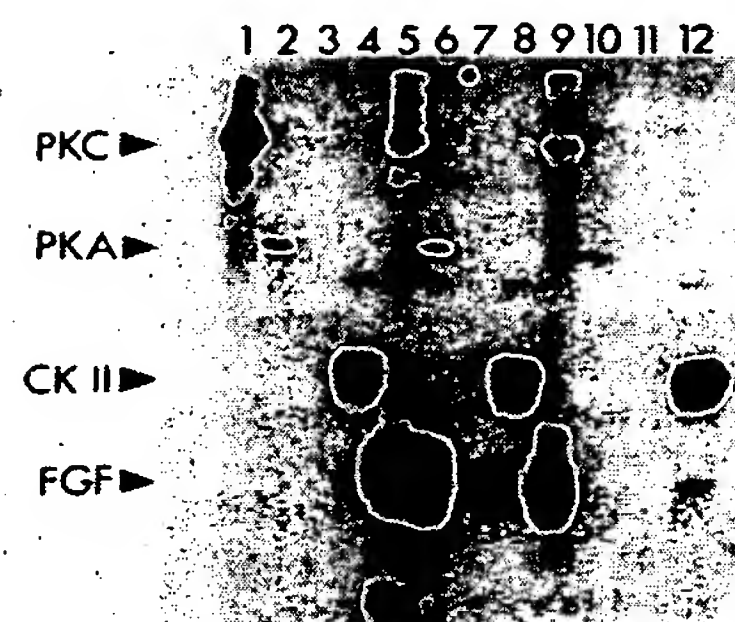


FIG. 1. Phosphorylation of recombinant acidic FGF and basic FGF by various purified protein kinases: Enzymes alone (lanes 1–4) or with 0.5  $\mu$ g of basic (lanes 5–8) or acidic (lanes 9–12) FGF were tested for their protein kinase activity. Purified PK-C (lanes 1, 5, and 9), PK-A (lanes 2, 6, and 10), casein kinase I (lanes 3, 7, and 11) or casein kinase II (CK II) (lanes 4, 8, and 12) were incubated under the standard phosphorylation conditions described in *Materials and Methods*, and the radiolabeled proteins were separated by SDS/15% PAGE and visualized by autoradiography. Arrowheads show the position of the autophosphorylated enzymes and of the phosphorylated growth factor.

phosphorylation was also shown to be strongly phospholipid dependent (not shown). Similar results shown in lane 6 of Fig. 1 establish that human basic FGF is a good substrate for PK-A as well. There was no evidence that basic FGF is a substrate for casein kinase I (lane 7), and only trace amounts of the mitogen were phosphorylated by casein kinase II (lane 8). An analysis of the results with acidic FGF are presented in lanes 9–12. While human acidic FGF was a substrate for PK-C-dependent phosphorylation (lane 9), PK-A and casein kinase I were ineffective (lanes 10 and 11), and casein kinase II had a small but negligible effect (lane 12).

**Identification of the Site of Phosphorylation.** Phosphoamino acid analysis of radiolabeled human basic FGF revealed the presence of phosphoserine after incubation with PK-C and phosphothreonine after incubation with PK-A (Fig. 2). Thus, the amino acid target sites phosphorylated by PK-C and

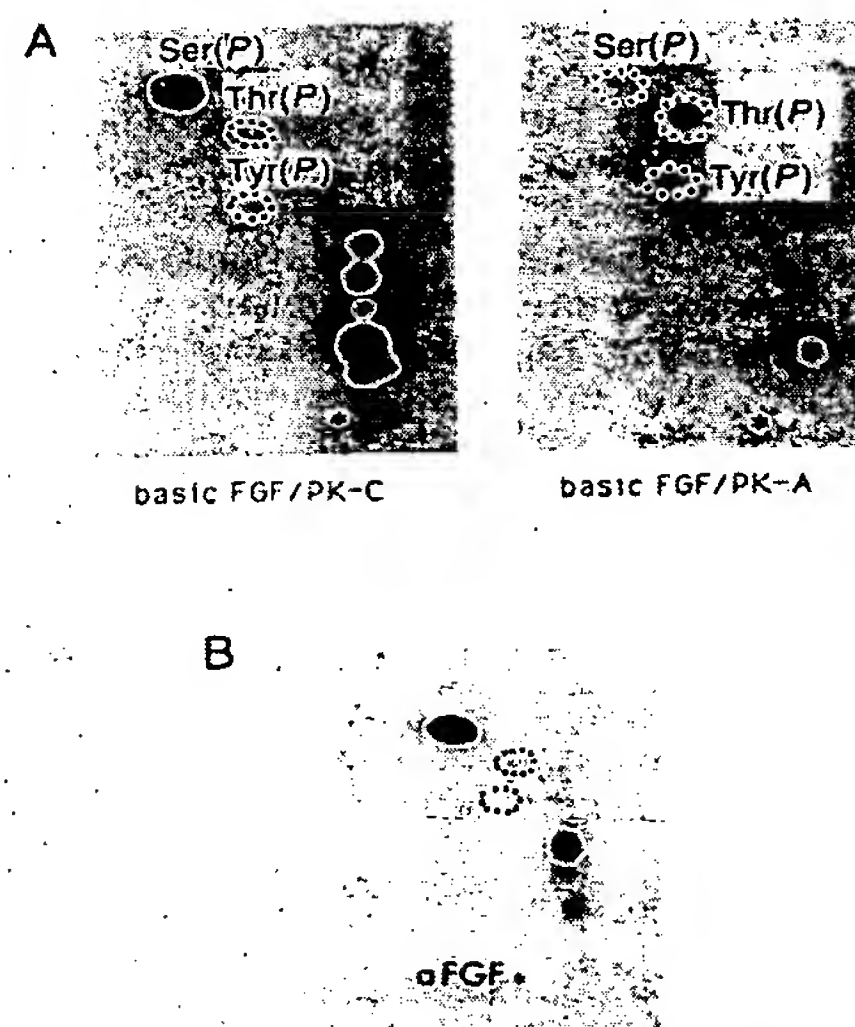


FIG. 2. Identification of the amino acids targeted by PK-C and PK-A. Recombinant human basic FGF was phosphorylated by PK-C or PK-A, and recombinant acidic FGF was phosphorylated by PK-C. Radiolabeled amino acids were identified by two-dimensional electrophoresis after partial hydrolysis in 6 M HCl. PK-A has no effect on acidic FGF and was not analyzed.



PK-A in human basic FGF are distinct. Three sites in the sequence of basic FGF are compatible with the consensus sequence (27) required for PK-C-dependent phosphorylation: Ser-64, Ser-108, and Ser-143. Reverse-phase HPLC analysis of the fragments generated by *Staphylococcus* V<sub>8</sub> protease digestion of the basic FGF phosphorylated by PK-C identified Ser-64 as the target amino acid for this enzyme (J.-J.F., unpublished data).

The primary structure of human basic FGF contains two potential sites that meet the criteria for PK-A-dependent phosphorylation (Thr-112 and Ser-113). Because threonine was identified as the target amino acid for the PK-A-dependent phosphorylation (Fig. 2), it is possible to propose Thr-112 as the targeted amino acid for PK-A. This is further supported by the observation that bovine basic FGF, in which Thr-112 has replaced Ser-112, was phosphorylated by PK-A on a target serine rather than threonine (result not shown). It is particularly interesting to note that this PK-A site (Thr-112) is located in the receptor-binding domain of basic FGF (28) and that all forms of basic FGF (human, bovine, ovine, rat, and frog) contain this potential site of phosphorylation. This is in contrast to the other members of the FGF family including acidic FGF (2) and the proteins encoded by the oncogenes *int-2* (29), *FGF-5* (30) and *hst/ks* (31, 32).

**Specificity of FGF Phosphorylation.** Several protein kinases were tested for their ability to phosphorylate acidic and basic FGFs in an effort to establish the possible specificity of PK-A and PK-C. Casein kinases I and II, enzymes that usually prefer acidic substrates (18, 21), did not significantly catalyze the phosphorylation of either mitogen (Fig. 1). The tyrosine kinase of the epidermal growth factor (EGF) receptor was also unable to phosphorylate basic FGF (C. Cochet, personal communication), although gag-fps, an oncogenic fusion protein with broad protein tyrosine kinase activity (provided to us by K. Gould, Salk Institute), was able to phosphorylate basic FGF and, to a larger extent, acidic FGF (not shown).

**Kinetics of Phosphorylation.** The  $K_m$  and  $V_{max}$  values were deduced from Lineweaver-Burke analyses of the reaction, and are shown in Table 1. The  $K_m$  values (1.5 and 3.4  $\mu M$ ) measured for the PK-C-dependent phosphorylation of basic and acidic FGFs rank these two factors among the better substrates for this enzyme (27). As an example, EGFR1, a peptide analog of the site of PK-C-dependent phosphorylation in the EGF receptor, has a  $K_m$  of 15  $\mu M$  (21). The  $V_{max}$  for the phosphorylation of basic FGF was 50 times greater for basic FGF than for acidic FGF, suggesting that it can be more extensively phosphorylated. The stoichiometry of phosphate incorporation was 0.5–0.6 mol of phosphate per mol of basic FGF, suggesting one site of phosphorylation.

The  $K_m$  value obtained for the phosphorylation of basic FGF by PK-A (17  $\mu M$ ) ranks basic FGF among the better substrates for this kinase (33). Although the rate ( $V_{max}$ , 0.06

nmol/min per mg) of this reaction is relatively low, 0.8–1.2 mol of phosphate are incorporated per mol of basic FGF. Acidic FGF is not a substrate for this enzyme.

**Basic FGF Exists as a Phosphorylated Protein in Cells Grown in Culture.** The hepatoma cell line SK-HEP-1 and bovine ACE cells were selected for analyses because they synthesize basic FGF (7, 8, 34). Detergent extracts of the phosphate-labeled cells were immunoprecipitated with a specific anti-basic FGF antibody (35, 36), and two major radiolabeled bands were detected by autoradiography after electrophoresis (Fig. 3). In each instance, one of the bands could be identified as being basic FGF by virtue of antibody specificity, the estimated molecular weights of the radiolabeled bands (16–18 kDa), and their displacement by the addition of unlabeled FGF during the immunoprecipitation. The phosphorylated basic FGF synthesized by ACE cells was further characterized by showing that it could be extracted with 2 M NaCl and eluted with a characteristic 1.4–1.6 M NaCl from the heparin-Sepharose column (not shown). No radiolabeled basic FGF was detected in conditioned media.

**Effects of Phosphorylation on the Biological Activities of Basic FGF.** No differences were found when PK-C-phosphorylated basic FGF was tested on capillary (Fig. 4A) or vascular (not shown) endothelial cell proliferation assays. There also were no effects on thymidine incorporation into 3T3 cells (not shown). Because each of these assays are long-term assays consisting of incubations of 24–96 hr, this basic FGF was also tested in a radioreceptor assay (Fig. 4B). There were no differences between this form of phosphorylated basic FGF and the recombinant material as determined by their capacity to displace labeled basic FGF bound to either high- or low-affinity sites on BHK cells. There were also no effects on the binding of basic FGF to immobilized heparin (not shown).

It was of particular interest to determine the effects of PK-A-phosphorylated FGF because its target amino acid (Thr-112) is located in the receptor-binding domain of basic FGF (28). As expected, the time course and experimental conditions of the proliferation (Fig. 4C) and mitogenesis (not shown) assays precluded any detection of difference between the unphosphorylated and PK-A-phosphorylated basic FGF. In contrast, the phosphorylated FGF was 3–8 times more potent at displacing <sup>125</sup>I-labeled FGF binding to its receptor than the unphosphorylated recombinant basic FGF. The effect of PK-A-phosphorylated FGF was shown to be specific for the high-affinity receptor because the displacement of FGF bound to cell-associated glycosaminoglycans was not different between the phosphorylated and unphosphorylated

Table 1. Kinetic parameters of the phosphorylation of FGFs by PK-C and PK-A

Protein kinase	Substrate	$K_m$ , $\mu M$	$V_{max}$ , nmol/min per mg
PK-C	bFGF	1.5	18
	aFGF	3.4	0.37
PK-A	bFGF	17.0	0.06
	aFGF	NS	NS

Various amounts of human recombinant acidic FGF (aFGF) and basic FGF (bFGF) were phosphorylated for 15 min at 30°C in the presence of [ $\gamma$ -<sup>32</sup>P]ATP (10  $\mu M$ ; 1500 cpm/pmol) and PK-C or PK-A under the conditions described in Fig. 1. Radioactivity incorporated into the trichloroacetic acid-precipitable proteins was determined and corrected for the autophosphorylation of the kinase. The kinetic parameters were calculated from Lineweaver-Burke double-reciprocal plots of the results. NS, not a substrate.

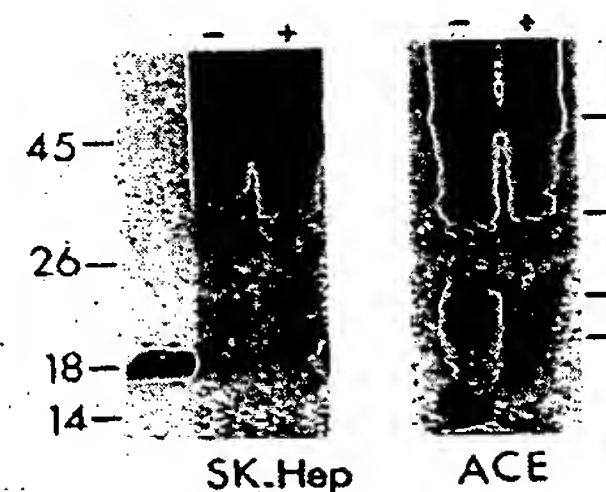


FIG. 3. Phosphorylation of basic FGF in cultured bovine ACE cells and human hepatoma SK-HEP-1 cells. Immunoprecipitates from <sup>32</sup>P-labeled cells were prepared as discussed, analyzed by SDS/15% PAGE, and visualized by autoradiography. Immunoprecipitations were performed in the absence (–) or in the presence (+) of an excess of nonradioactive basic FGF. The position of the molecular weight standards ( $\times 10^{-3}$ ) is shown on both sides of the figure, and radioiodinated recombinant FGF is shown in the far-left lane.



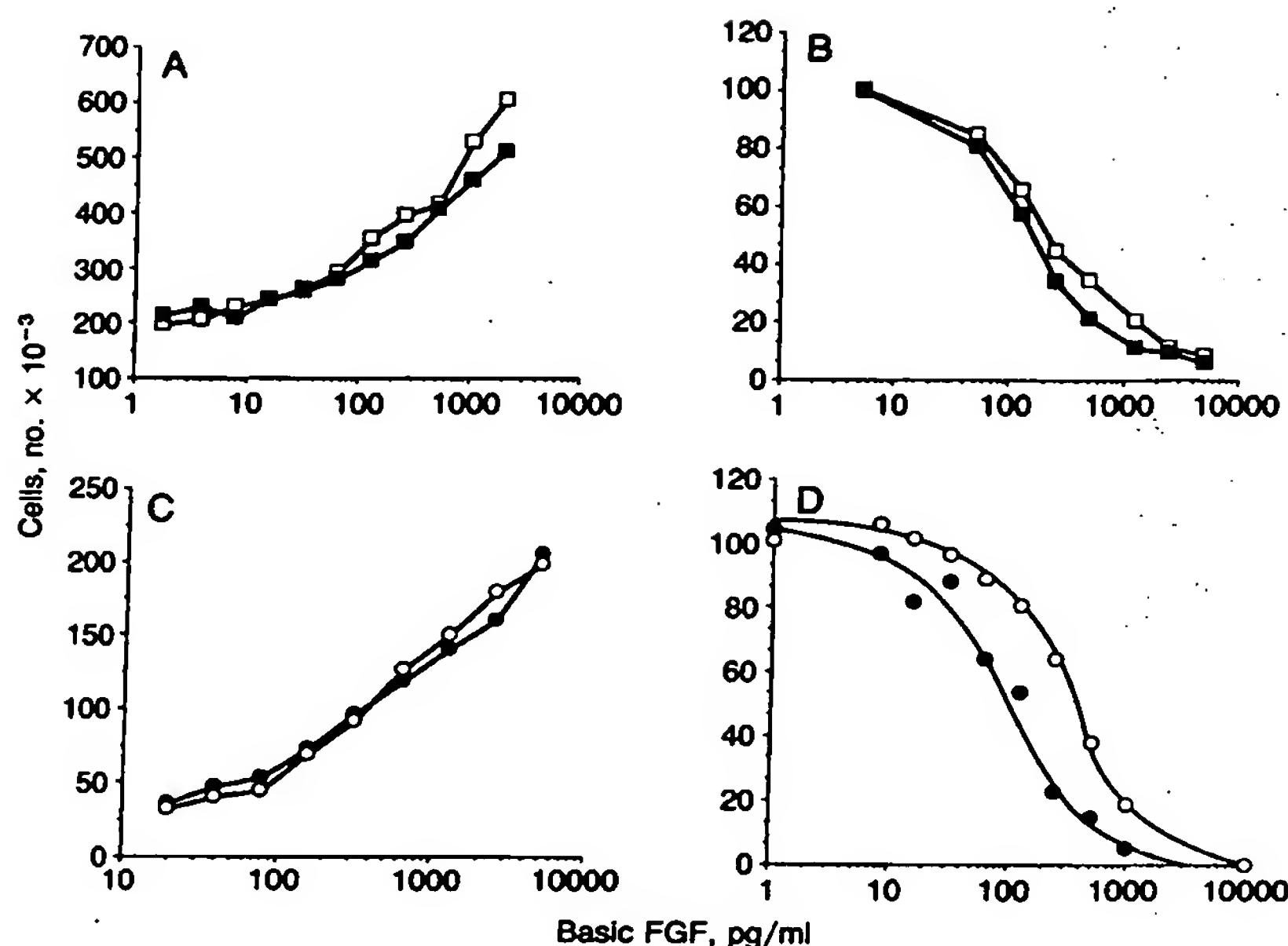


FIG. 4. Biological activities of phosphorylated basic FGFs. Recombinant human basic FGF was phosphorylated (■) by PK-C and tested for its capacity to stimulate the proliferation of ACE cells (A) and for its ability to displace <sup>125</sup>I-labeled FGF binding to its receptor in BHK cells (B). Basic FGF was phosphorylated by PK-A (●) and tested for its capacity to stimulate the proliferation of ACE cells (C) and to displace <sup>125</sup>I-labeled FGF binding to the FGF receptor on BHK cells (D). In each instance, control samples (□, ○) were treated under identical conditions, but in the absence of enzyme.

basic FGFs. Accordingly, the affinity of the phosphorylated FGF for immobilized heparin was also found to be unchanged by phosphorylation (not shown). We also exploited the fact that the site of phosphorylation of basic FGF by PK-A can be modified by heparin so that the enzyme phosphorylates Ser-64 rather than Thr-112 (37). Under these conditions, there were no differences in any of the assays between the phosphorylated and unphosphorylated forms of basic FGF (data not shown).

**Growth Factors Are Substrates for Phosphorylation.** TGF- $\alpha$ , TGF- $\beta$ , and IGF-I were substrates for PK-C-mediated phosphorylation (Fig. 5A). Not all growth factors were substrates; however, EGF, PDGF, and NGF were either poorly or not-at-all phosphorylated by this kinase. In the case of PK-A, TGF- $\alpha$  and TNF were substrates for phosphorylation, but EGF, PDGF, and NGF were not. The spectrum of

growth factors phosphorylated by different kinases varied presumably by virtue of the presence of consensus phosphorylation sequences. Structurally related factors exhibited very different substrate specificities. Basic FGF and TGF- $\alpha$  were phosphorylated by both PK-C and PK-A, yet the structural homolog acidic FGF, which has 55% structural identity with basic FGF, was not phosphorylated by PK-A, and EGF, which is 33% homologous to TGF- $\alpha$ , was not phosphorylated by either kinase.

## DISCUSSION

It is not known at what step in its synthesis basic FGF is phosphorylated or in fact what enzymes are responsible for the phosphorylation of basic FGF in intact cells. Attempts to identify the amino acid phosphorylated *in vivo* have so far failed because of the small amounts of growth factor made by the cells (7–10) and the low recoveries from immunoprecipitation, peptide hydrolysis, and two-dimensional electrophoresis. The availability of transfected cells with a high expression of basic FGF (11, 12, 38) should alleviate this problem.

Little, if anything, is known about the mechanisms that regulate synthesis, secretion, and bioavailability of growth factors, and the results presented here suggest that it will be important to consider the possible roles of phosphorylation. In the case of basic FGF, one such role includes targeting to specific subcellular organelles. The metabolism of basic FGF after binding to its receptor is different from many other growth factors (39). It is specifically metabolized to three long-lived fragments that continue to retain their capacity to bind heparin. Bouché *et al.* (40) have suggested that it is targeted to and accumulates in the nucleolus after binding to its receptor on the plasma membrane. Whether this transport is kinase dependent is not known. With the observation that FGFs have no obvious signal sequence, remain cell-associated, and may not be secreted in a classical sense (1, 2, 7), the targeting of these mitogens to specific locations for processing in the cell may be of particular importance.

The observation that the PK-A-phosphorylated basic FGF has a greater capacity to displace FGF from its receptor supports the notion that it possesses a higher affinity for the FGF receptor. If this is the case, then this posttranslational change may well contribute to the regulation of FGF activity.

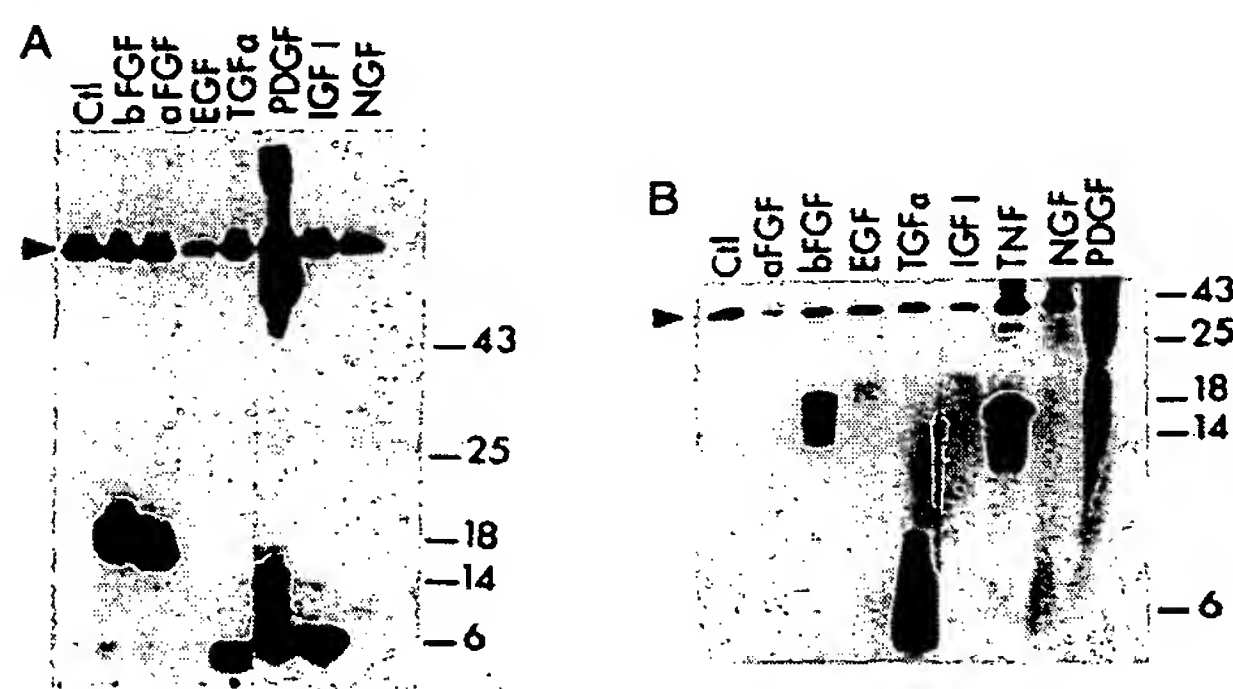


FIG. 5. Phosphorylation of growth factors by purified PK-C (A) and PK-A (B). The growth factors indicated (0.5  $\mu$ g per lane) were phosphorylated as described in the text. The reactions were stopped by the addition of Laemmli sample buffer, and the phosphorylated proteins were visualized by autoradiography after separation by SDS/15% PAGE (A) or SDS/20% PAGE (B). Positions of the molecular weight markers (shown  $\times 10^{-3}$ ) are indicated in the far-right lane. Arrowheads indicate the position of the autophosphorylated protein kinases. Because PDGF was stored in the presence of bovine serum albumin, only 0.1  $\mu$ g of this factor was used, and the phosphorylation of bovine serum albumin can be detected as a 67-kDa band. Ctl, control; bFGF, basic FGF; aFGF, acidic FGF.

Increased binding of FGF to its receptor has recently been proposed as one of the mechanisms that might enable cells to recruit FGF from the extracellular matrix. Thus, it is interesting to speculate that protein kinases mediate the release of FGFs from the basement membranes. This notion is supported by the work of Plouët *et al.* (17), who recently have demonstrated that the release of acidic FGF from retinal rod outer segment membranes requires ATP-dependent phosphorylation. Any involvement of protein kinases would invoke the existence of these enzymes either in the extracellular milieu or associated with the outer plasma membrane.

The results presented here may have several important implications for our understanding of the mechanisms regulating growth factor activity. Because basic FGF, and possibly other growth factors, are phosphorylated in intact cells, their biological activities, bioavailability, and/or specificities may well be different from the unphosphorylated forms currently being examined in experimental models. The failure to establish that the phosphorylation of basic FGF results in modified activities in the proliferation and mitogenic assays most likely reflects the lability of this posttranslational modification. This supports a potential role for ectokinases and ectophosphatases in regulating the biological activities of basic FGF. The balance between kinases and phosphatases would effectively force a local action for phosphorylated basic FGF. Because the site of phosphorylation of basic FGF by PK-A is modified by heparin (37) and this site of phosphorylation is required for increased receptor binding, protein kinase action is ineffective at modifying FGF activity when it is bound to glycosaminoglycans in the extracellular matrix. The generation of basic FGF free of any heparin or related glycosaminoglycans will be prerequisite to observing any effect of PK-A on basic FGF *in vivo* and, thus, supports the potential role of heparinases in regulating FGF activity (8).

It is important to note that the experimental parameters examined here represent only three (binding, mitogenesis, and proliferation) effects of basic FGF. It will be of importance to establish whether the phosphorylation of FGF modulates cell differentiation (41) and cell adhesion (42). The phosphorylation of potentially cytoplasmic and certainly cell-associated growth factors like the FGFs may well provide a critical regulatory step on signal transduction and any one of their many diverse biological end points.

We are grateful for the expert technical assistance of Jim Farris and Emelie Amburn and for the skillful secretarial work of Denise Higgins and to Drs. D. Schubert, T. Hunter, and I. Vilgrain for critical reading of the manuscript. This research was supported by grants from the National Institutes of Health (HD-09690 and DK-18811), from the Robert J. Kleberg and Helen C. Kleberg Foundation, and from the G. Harold and Leila Y. Mathers Charitable Foundation. J.-J.F. was a visiting scientist from Institut National de la Santé et de la Recherche Médicale (Unité 244, Grenoble).

- Esch, F., Baird, A., Ling, N., Ueno, N., Hill, F., Denoroy, L., Klepper, R., Gospodarowicz, D., Böhlen, P. & Guillemin, R. (1985) *Proc. Natl. Acad. Sci. USA* 82, 6507-6511.
- Gimenez-Gallego, G., Rodkey, K., Bennett, C., Rios-Candelore, M., DiSalvo, J. & Thomas, K. A. (1985) *Science* 230, 1385-1387.
- Jaye, M., Howk, R., Burgess, W., Ricca, G. A., Chiu, I. M., Raver, M. W., O'Brien, S. J., Modi, W. S., Maciag, T. & Drohan, W. N. (1986) *Science* 233, 541-545.
- Abraham, J. A., Mergia, A., Whang, J. L., Tumolo, A., Friedman, J., Hjerrield, K. A., Gospodarowicz, D. & Fiddes, J. C. (1986) *Science* 233, 545-548.
- Neufeld, G. & Gospodarowicz, D. (1985) *J. Biol. Chem.* 260, 13860-13868.
- Friesel, R., Burgess, W. H., Mehlman, T. & Maciag, T. (1986) *J. Biol. Chem.* 261, 7581-7584.
- Vlodavsky, I., Fridman, R., Sullivan, R., Sasse, J. & Klagsbrun, M. (1987) *J. Cell. Physiol.* 131, 402-408.
- Baird, A. & Ling, N. (1987) *Biochem. Biophys. Res. Commun.* 142, 428-435.
- Vlodavsky, I., Folkman, J., Sullivan, R., Fridman, R., Ishai-Michaeli, R., Sasse, J. & Klagsbrun, M. (1987) *Proc. Natl. Acad. Sci. USA* 84, 2292-2296.
- Folkman, J., Klagsbrun, M., Sasse, J., Wadzinski, M. G., Ingber, D. & Vlodavsky, I. (1988) *Am. J. Pathol.* 130, 393-400.
- Rogelj, S., Weinberg, R. A., Fanning, P. & Klagsbrun, M. (1988) *Nature (London)* 331, 173-175.
- Jaye, M., Lyall, R. M., Mudd, R., Schlessinger, J. & Sarver, N. (1988) *EMBO J.* 7, 963-969.
- Hunter, T. & Cooper, J. A. (1985) *Annu. Rev. Biochem.* 54, 897-930.
- Feige, J. J. & Chambaz, E. M. (1987) *Biochimie* 69, 379-385.
- Nishizuka, Y. (1984) *Nature (London)* 308, 693-698.
- Berridge, M. J. & Irvine, R. F. (1984) *Nature (London)* 312, 315-321.
- Plouët, J., Mascarelli, F., Loret, M. D., Faure, J. P. & Courtois, Y. (1988) *EMBO J.* 7, 373-376.
- Cochet, C., Job, D., Pirollet, F. & Chambaz, E. M. (1980) *Endocrinology* 106, 750-757.
- Cochet, C., Feige, J. J. & Chambaz, E. M. (1983) *Biochim. Biophys. Acta* 743, 1-12.
- Barr, P. J., Cousens, L. S., Lee-Ng, C. T., Medina-Selby, A., Masiarz, F. R., Hallewell, R. A., Chamberlain, S., Bradley, J., Lee, D., Steimer, K. S., Poulter, L., Burlingame, A. L., Esch, F. & Baird, A. (1988) *J. Biol. Chem.* 263, 16471-16478.
- Walton, G. M., Bertics, P. J., Hudson, L. G., Vedvick, T. S. & Gill, G. N. (1987) *Anal. Biochem.* 161, 425-437.
- Laemmli, U. K. (1970) *Nature (London)* 227, 680-685.
- Cooper, J. A., Sefton, B. M. & Hunter, T. (1983) *Methods Enzymol.* 99, 387-402.
- Sefton, B. M., Beemon, K. & Hunter, T. J. (1978) *J. Virol.* 28, 957-971.
- Moscattelli, D. (1987) *J. Cell Physiol.* 131, 123-130.
- Gospodarowicz, D., Massaglia, S., Cheng, J. & Fujii, D. K. (1986) *J. Cell. Physiol.* 127, 121-136.
- Woodgett, J. R., Gould, K. L. & Hunter, T. (1986) *Eur. J. Biochem.* 161, 177-184.
- Baird, A., Schubert, D., Ling, N. & Guillemin, R. (1988) *Proc. Natl. Acad. Sci. USA* 85, 2324-2328.
- Dickson, C. & Peters, G. (1987) *Nature (London)* 326, 833.
- Zhan, X., Bates, B., Hu, X. & Goldfarb, M. (1988) *Mol. Cell. Biol.* 8, 3487-3497.
- Yoshida, T., Miyagawa, K., Odagiri, H., Sakamoto, H., Little, P. F., Terada, M. & Sugimura, T. (1987) *Proc. Natl. Acad. Sci. USA* 84, 7305-7309.
- Delli Bovi, P., Curatola, A. N., Kern, F. G., Greco, A., Ittman, M. & Basilico, C. (1987) *Cell* 50, 729-737.
- Carlson, G. M., Bechtel, P. J. & Graves, D. J. (1979) *Adv. Enzymol.* 50, 41-115.
- Schweigerer, L., Neufeld, G., Friedman, J., Abraham, J. A., Fiddes, J. C. & Gospodarowicz, D. (1987) *Nature (London)* 325, 257-259.
- Halaban, R., Ghosh, S. & Baird, A. (1987) *In Vitro Cell. Dev. Biol.* 23, 47-52.
- Halaban, R., Kwon, B., Ghosh, S., Delli Bovi, P. & Baird, A. (1988) *Mol. Cell. Biol.* 8, 2933-2941.
- Feige, J. J. & Baird, A. (1989) *J. Cell. Biochem. Suppl.* 13B, 154 (abstr.).
- Thomas, K. A. (1988) *Trends Biochem. Sci.* 13, 327-328.
- Moenner, M., Badet, J., Chevallier, B., Tardieu, M., Courty, J. & Barritault, D. (1987) in *Angiogenesis: Mechanisms and Pathobiology*, eds. Rifkin, D. B. & Klagsbrun, M. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 52-57.
- Bouché, G., Gas, N., Prats, H., Balsin, V., Tauber, J. P., Teissié, J. & Amalric, F. (1987) *Proc. Natl. Acad. Sci. USA* 84, 6770-6774.
- Gospodarowicz, D., Neufeld, G. & Schweigerer, L. (1987) *Endocrinol. Rev.* 8, 95-114.
- Schubert, D., Ling, N. & Baird, A. (1987) *J. Cell Biol.* 104, 635-643.



# A variant epidermal growth factor receptor exhibits altered type $\alpha$ transforming growth factor binding and transmembrane signaling

(protooncogenes)

TETSUYA MORIAI\*, MICHAEL S. KOBRIN, CHRISTOPHER HOPE, LYNN SPECK, AND MURRAY KORC†

Departments of Medicine and Biological Chemistry, University of California, Irvine, CA 92717

Communicated by Morris E. Friedkin, July 20, 1994

**ABSTRACT** Epidermal growth factor (EGF) and type  $\alpha$  transforming growth factor (TGF- $\alpha$ ) bind to a specific region in subdomain III of the extracellular portion of the EGF receptor (EGFR). Binding leads to receptor dimerization, auto- and transphosphorylation on intracellular tyrosine residues, and activation of signal transduction pathways. We compared the binding and biological actions of EGF and TGF- $\alpha$  in Chinese hamster ovary (CHO) cells expressing either wild-type human EGFR (HER497R) or a variant EGFR that has an arginine-to-lysine substitution in the extracellular domain at codon 497 (HER497K) within subdomain IV of EGFR. Both receptors exhibited two orders of binding sites with radiolabeled EGF ( $^{125}$ I-EGF). Similar results were obtained with  $^{125}$ I-TGF- $\alpha$  in cells expressing HER497R. In contrast, only one order of low-affinity binding sites was seen with  $^{125}$ I-TGF- $\alpha$  in the case of HER497K. Although EGF and TGF- $\alpha$  enhanced tyrosine phosphorylation of both receptors, CHO cells expressing HER497K exhibited an attenuated growth response to EGF and TGF- $\alpha$  and a reduced induction of the protooncogenes *FOS*, *JUN*, and *MYC*. Moreover, high concentrations of TGF- $\alpha$  (5 nM) inhibited growth in these cells but not in cells expressing HER497R. These findings indicate that a region in subdomain IV of EGFR regulates signal transduction across the cell membrane and selectively modulates the binding characteristics of TGF- $\alpha$ .

The family of transmembrane tyrosine kinase receptors constitutes an important group of regulatory proteins, which include the insulin receptor and the epidermal growth factor receptor (EGFR). Ligand specificity is conferred by the presence of unique sequences in the extracellular domain of each receptor. Often, more than one ligand can bind to the same receptor. Thus, the insulin receptor binds insulin and, with a lesser affinity, insulin-like growth factors I and II (1). EGFR binds EGF, type  $\alpha$  transforming growth factor (TGF- $\alpha$ ), amphiregulin, betacellulin, and heparin-binding EGF-like growth factor (2, 3). Ligand binding in the extracellular domain leads to receptor autophosphorylation on tyrosine residues located within the intracellular domain and initiation of a cascade of biochemical reactions that mediate the biological actions of many hormones and growth factors (1-3). In the absence of ligand, the extracellular domain acts through unknown mechanisms to prevent constitutive and unregulated activation of the intrinsic tyrosine kinase activity of the intracellular domain (4, 5).

The extracellular domain of transmembrane tyrosine kinase receptors often contains different structural subdomains. In the case of EGFR, there are four contiguous regions consisting of subdomain I at the N terminus and the cysteine-rich subdomains II and IV flanking subdomain III (6). Cross-linking studies have identified a 47-amino acid

sequence (residues 321-367) within subdomain III as the EGF/TGF- $\alpha$  binding region (6, 7). The specific functions of the other extracellular subdomains are not known.

Recently, a variant EGFR has been identified that has an arginine-to-lysine substitution at codon 497 (8). This variant receptor was found in normal human lymphocytes and a number of cultured human cancer cell lines but not in A431 cells (8). Inasmuch as the potential role of this region of EGFR with respect to modulation of cell proliferation and gene activation is not known, in the present study we sought to characterize the biological properties of HER497K. Accordingly, we expressed the wild-type human EGFR (HER497R) and HER497K in Chinese hamster ovary (CHO) cells devoid of endogenous EGFR and determined whether there are differences between the two receptors with respect to ligand binding, cell growth, and induction of the immediate early response genes *FOS*, *JUN*, and *MYC*.

## MATERIALS AND METHODS

**Materials.** The following materials were purchased: receptor grade murine EGF from Collaborative Research; GF/F filters from Whatman; PY-20 anti-phosphotyrosine antibodies from ICN;  $^{125}$ I, [ $\alpha$ - $^{32}$ P]dCTP, ECL blotting kit, and Hyperfilm-ECL from Amersham; Iodo-Gen from Pierce; pGEM-7Zf(+) plasmid from Promega; pSVK 3 expression vector from Pharmacia; human *MYC*, *FOS*, and *JUN* cDNAs from the American Type Culture Collection; GeneScreen membranes from New England Nuclear; Qiagen minicolumns; 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) from Sigma; and Ecocint scintillation fluid from National Diagnostics. Recombinant human TGF- $\alpha$ , 13A9 monoclonal anti-EGFR antibodies (9), and CHO cells were gifts from M. Winkler and B. Fendley (Genentech, South San Francisco).

**Generation of Expression Vectors.** Two overlapping oligonucleotides were synthesized (8) and used to generate a linker that contained the following restriction sites: *Sph* I/*Hind* III/*Sac* I/*Eco* RI/*Bgl* II/*Sma* I/*Sal* I/*Bst* XI. The linker was subcloned into the cloning vector pGEM-7Zf at the *Sph* I/*Bst* XI sites, generating the vector pGEM7.2Zf. Four *Eco* RI fragments (HERA, HERB, HERC, and HERD) encompassing the entire EGFR coding region were isolated from a  $\lambda$  library derived from T<sub>3</sub>M<sub>4</sub> human pancreatic cancer cells, with HERB containing the nucleotide base change (8). The wild-type EGFR expression plasmid fragment was prepared by substituting HERB with fragment HERG, originally cloned from A431 cells (2, 25). The four EGFR fragments

Abbreviations: EGF, epidermal growth factor; EGFR, EGF receptor; TGF- $\alpha$ , type  $\alpha$  transforming growth factor; CHO, Chinese hamster ovary; MTT, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide.

\*Present address: Second Department of Internal Medicine, Asahikawa Medical College, Asahikawa, Japan.

†To whom reprint requests should be addressed.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.



were reassembled as follows: the 314-bp *Sac* I/*Eco*RI fragment of HERA, the 1841-bp *Eco*RI HERB (or HERG) fragment, and the 637-bp *Eco*RI/*Bgl* II fragment of HERC were sequentially subcloned into pGEM-7.2Zf, generating plasmid HER7.2ABC. The 771-bp *Eco*RI HERC fragment and the 884-bp *Eco*RI/*Sca* I fragment of HERD were sequentially subcloned into pGEM-7.2Zf, generating plasmid HER7.2CD. A 1025-bp *Bgl* II/*Sal* I fragment from HER7.2CD was then subcloned into HER7.2ABC, yielding a full-length EGFR cDNA. The two 3.8-kb *Hind*III/*Sal* I EGFR cDNAs were then subcloned into the pSVK 3 expression vector under the control of the simian virus 40 early promoter, yielding pHER497K (variant EGFR) and pHER497R (wild-type EGFR). Authenticity of the constructs was confirmed by sequencing. Transfection of pHER497K, pHER497R, and pSV2-*dhfr* DNA into CHO cells deficient in synthesis of dihydrofolate reductase was carried out by the calcium phosphate precipitation technique (10). After reaching confluency, cells were plated at a 1:10 dilution in selection medium [Dulbecco's modified Eagle's medium (DMEM)/Ham's F-12 medium (1:1) lacking hypoxanthine, glycine, and thymidine] supplemented with 5% dialyzed fetal bovine serum (FBS). After 2 weeks, single clones of cells were isolated, and each individual clone was plated separately. After clonal expansion, cells from each individual clone were tested for EGFR expression by Northern blot analysis and  $^{125}$ I-labeled EGF ( $^{125}$ I-EGF) binding. The selected clones were shown to express only the variant HER497K or wild-type HER497R by subjecting RNA samples to reverse transcriptase PCR amplification followed by asymmetric PCR and sequencing (8).

**Binding Experiments.** Biologically active  $^{125}$ I-EGF (40–50  $\mu$ Ci/ $\mu$ g; 1 Ci = 37 GBq) and  $^{125}$ I-TGF- $\alpha$  (50–70  $\mu$ Ci/ $\mu$ g) were iodinated with Iodo-Gen and chloramine T, respectively (11). Binding was performed at 4°C for 5 hr on cells in monolayer culture (70–80% confluent) with increasing concentrations of labeled ligand and DMEM containing 0.1% bovine serum albumin (BSA) and 20 mM Hepes (11). Incubations were stopped by washing cells with phosphate-buffered saline containing 0.1% BSA. Nonspecific binding, determined in the presence of a 1000-fold excess of unlabeled EGF, did not exceed 10% of total binding. Equilibrium binding data were analyzed with the LIGAND program (12).

**Cell Growth Assay.** Cell proliferation was assayed by measuring the metabolism of the tetrazolium salt MTT and by monitoring [ $^3$ H]thymidine incorporation into DNA (13). The MTT assay measures the cell's ability to metabolize MTT, which correlates well with the rate of cellular proliferation (13). For the MTT experiments, cells were plated at a density of  $8 \times 10^3$  cells per well in 96-well plates and grown overnight in 0.2 ml of DMEM/Ham's F-12 medium (1:1) containing 5% FBS. After replacing the medium with 0.2 ml of the same medium containing 0.5% FBS, cells were incubated for 72 hr at 37°C in the absence or presence of ligands, prior to the addition of MTT (5 mg/ml) for 4 hr (13). Reactions were stopped with acidified isopropanol (0.04 M), and the absorbance of the supernatants was measured at 570 nm with an ELISA plate reader (Molecular Devices). Absorbance readings were converted to the percentage of control growth, which was determined in the absence of growth factors and defined as 100% (13). For the [ $^3$ H]thymidine incorporation experiments, cells were plated at a density of  $3 \times 10^4$  cells per well in 12-well plates and incubated for 48 hr in 1.5 ml of DMEM/Ham's F-12 medium (1:1) containing 5% FBS. Cells were then incubated for 40 hr in the absence or presence of ligand with serum-free medium containing 5  $\mu$ g of insulin per ml, 5  $\mu$ g of transferrin per ml, 5 ng of selenium per ml, and 0.1% BSA before adding [ $^3$ H]thymidine (1  $\mu$ Ci/ml) for 1 hr. Cells were collected by rapid filtration on GF/F filters and washed sequentially with 10% trichloroacetic acid, 70%

ethanol, and 100% ethanol. Radioactivity was measured in a liquid scintillation counter using Ecoscint as the scintillant. Data were then expressed as percentage change from control values.

**EGFR Phosphorylation.** Cells were incubated for the specified times at 37°C in binding medium in the absence or presence of each ligand and then solubilized in buffer containing 50 mM Tris-HCl (pH 7.7), 150 mM NaCl, 5 mM EDTA, 1% deoxycholate, 1% Triton X-100, 0.1% sodium dodecyl sulfate, 1 mM phenylmethylsulfonyl fluoride, 2 mM benzamide, and 0.1 mM  $\text{Na}_3\text{VO}_4$  (14). Cell homogenates or immunoprecipitated EGFR (using the 13A9 anti-EGFR antibodies) were subjected to SDS/7.5% PAGE and transferred to nitrocellulose membranes (14). Membranes were incubated with PY-20 anti-phosphotyrosine antibodies (1  $\mu$ g/ml) for 1 hr, washed, and subjected to chemiluminescence Western blotting with the ECL blotting kit and Hyperfilm-ECL.

**Northern Blot Analysis.** Total RNA was isolated by the acid guanidium thiocyanate extraction method (15), size-fractionated, and transferred onto GeneScreen membranes (16). Filters were hybridized under high-stringency conditions using [ $\alpha$ - $^{32}$ P]dCTP-labeled cDNAs (16). The intensity of the radiographic bands was quantified by laser densitometry.

**Statistical Analysis.** Statistical analysis of the experimental results was obtained by one-tailed Student's *t* test using the STATVIEW computer program, with *P* < 0.05 taken as significant.

## RESULTS

**Characterization of Ligand Binding.** EGF and TGF- $\alpha$  binding and dissociation are pH dependent (14). Therefore, to determine EGFR affinity and capacity, equilibrium binding studies (4°C) were carried out at pH 7.4 (Fig. 1) and pH 8.0 (Table 1). Computer-based analysis of  $^{125}$ I-EGF binding data (12) demonstrated the presence of two orders of binding sites at the two pH levels in CHO cells expressing either wild-type HER497R or variant HER497K. At pH 7.4, there were approximately 90,000 and 116,000 low-affinity receptors in cells expressing HER497R and HER497K, respectively (Table 1). The number of high-affinity binding sites (3700 and 4800) was also similar in both cell types (Table 1). At pH 8.0, there was a decrease in the number of binding sites for EGF with both receptors and a concomitant increase in binding affinity (Table 1).  $^{125}$ I-TGF- $\alpha$  also exhibited two orders of binding sites in cells expressing HER497R. As in the case of EGF, there was a decrease in the number of TGF- $\alpha$  binding

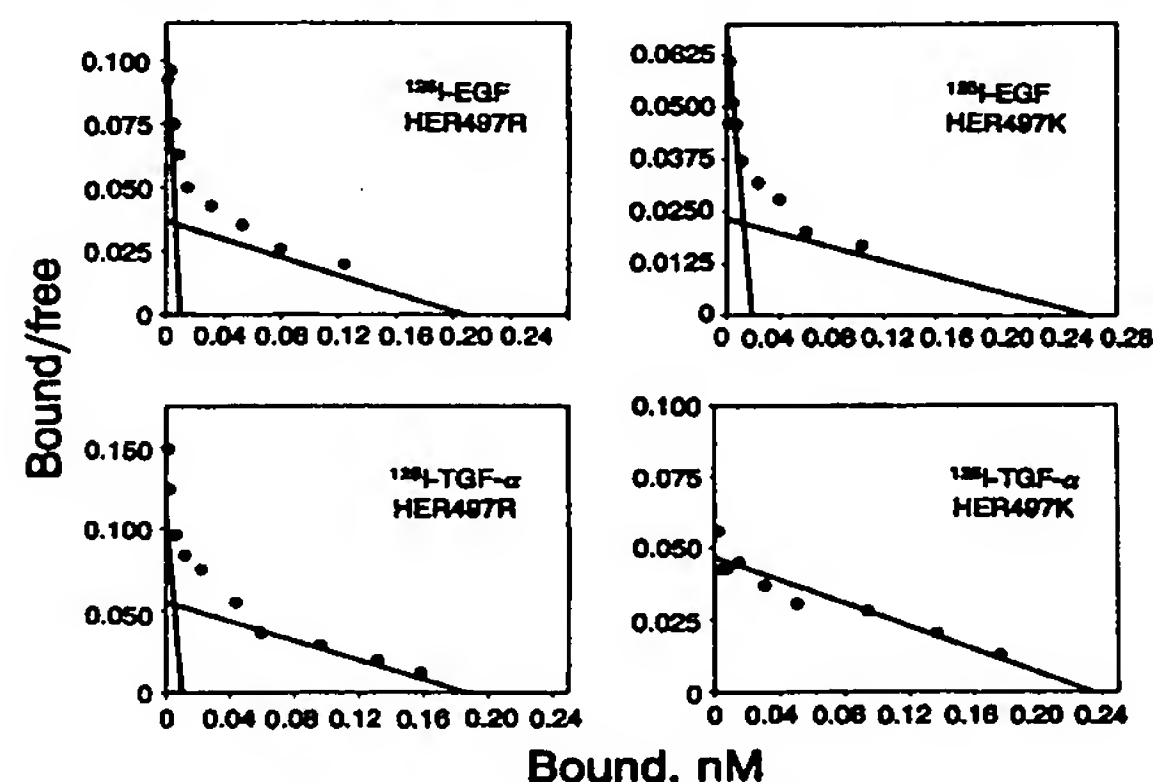


FIG. 1. Scatchard analysis. CHO cells expressing the indicated EGFR were incubated for 5 hr (4°C) in six-well plates with binding medium (pH 7.4) containing increasing concentrations (0.1–80 ng/ml) of either  $^{125}$ I-EGF or  $^{125}$ I-TGF- $\alpha$ . Data are representative of two experiments.

Table 1. Binding characteristics of  $^{125}\text{I}$ -EGF and  $^{125}\text{I}$ -TGF- $\alpha$ 

pH	Ligand	Receptor, no. per cell		$K_d$ , nM	
		HER497R	HER497K	HER497R	HER497K
7.4	EGF	89,900	116,200	5.73	13.6
		3,700	4,800	0.1	0.28
7.4	TGF- $\alpha$	123,300	158,100	4.57	5.62
		8,200		0.16	
8.0	EGF	47,300	83,800	3.39	8.49
		1,500	2,900	0.05	0.17
8.0	TGF- $\alpha$	61,600	91,400	2.26	3.75
		1,100		0.02	

HER497R- and HER497K-expressing cells were incubated for 5 hr (4°C) at the indicated pH in six-well plates with increasing concentrations of either  $^{125}\text{I}$ -EGF or  $^{125}\text{I}$ -TGF- $\alpha$  (0.1–80 ng/ml). Data were analyzed with the LIGAND program as described and are means of two separate experiments with each ligand.

sites at pH 8.0 in conjunction with an increase in binding affinity (Table 1). However, irrespective of the pH, only a single order of low-affinity binding sites was observed with  $^{125}\text{I}$ -TGF- $\alpha$  in cells expressing HER497K (Table 1).

To determine the effect of pH on ligand dissociation, cells were initially incubated for 5 hr at 4°C (pH 7.4) with either  $^{125}\text{I}$ -TGF- $\alpha$  or  $^{125}\text{I}$ -EGF, washed at 4°C, and incubated for 2 hr at various pH levels (4°C) in fresh medium devoid of radioactivity (Fig. 2). At pH 4, the majority of  $^{125}\text{I}$ -EGF was released from the cells. Raising the pH caused a progressive decrease in the percentage of dissociated radioactivity, with 50% dissociation occurring at pH 4.8 and 5.4 with HER497K and HER497R, respectively. The least dissociation occurred at pH 6.5–7.0, and the amount of  $^{125}\text{I}$ -EGF dissociated was similar with HER497R and HER497K. In contrast, with both receptors, the majority of  $^{125}\text{I}$ -TGF- $\alpha$  was still dissociated from the cells at pH 5, and 50% dissociation occurred at pH  $\approx$  7.0. However, at pH 8 and 8.5,  $^{125}\text{I}$ -TGF- $\alpha$  dissociated more readily from HER497K than from HER497R.

**Growth Stimulation by EGF and TGF- $\alpha$ .** As determined by the MTT dye reduction assay, EGF and TGF- $\alpha$  modulated

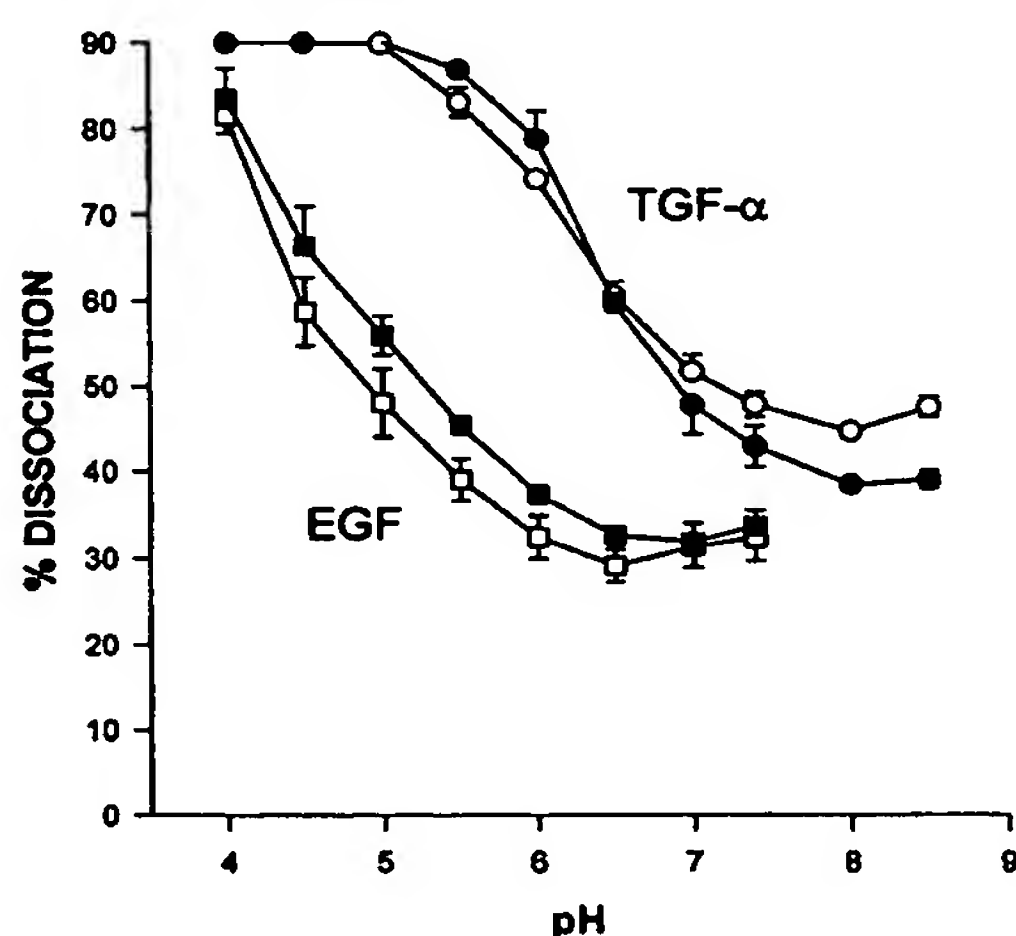


FIG. 2. Effect of pH on ligand dissociation. HER497R-expressing (solid symbols) and HER497K-expressing (open symbols) cells were incubated for 5 hr (4°C) in six-well plates with binding medium (pH 7.4) containing 200,000 cpm of  $^{125}\text{I}$ -EGF (squares) or  $^{125}\text{I}$ -TGF- $\alpha$  (circles) and then washed and incubated for 2 hr at 4°C in fresh medium at the indicated pH. Radioactivity dissociated from the cells during the 2-hr incubation was expressed as a percentage of the radioactivity that bound to the cells during the initial 5-hr incubation. Data are means  $\pm$  SE of duplicate determinations from three experiments.

the growth of EGFR-expressing CHO cells. In the case of HER497R, maximal stimulation of 40% occurred at concentrations of 0.1–0.8 nM with either growth factor (Fig. 3 A and C). Higher ligand concentrations resulted in a slight decrease in growth stimulation. In the case of HER497K, there was only a minimal increase in cell growth with EGF (0.1 nM) and none with TGF- $\alpha$  (Fig. 3 A and C). Furthermore, at high concentrations (2–5 nM), TGF- $\alpha$  inhibited cell proliferation (Fig. 3C). When [ $^3\text{H}$ ]thymidine incorporation was assayed, maximal stimulation of 45–90% occurred at concentrations of 0.2–2 nM EGF and TGF- $\alpha$  with HER497R (Fig. 3B). In contrast, with HER497K, low concentrations of EGF increased [ $^3\text{H}$ ]thymidine incorporation by 45%, whereas TGF- $\alpha$  was without effect. Conversely, high concentrations of TGF- $\alpha$  (1.5–5 nM) markedly inhibited [ $^3\text{H}$ ]thymidine incorporation, and EGF was without effect (Fig. 3D).

**Effects of EGF and TGF- $\alpha$  on EGFR Phosphorylation.** To determine whether EGF and TGF- $\alpha$  induce tyrosine phosphorylation of HER497K, immunoblotting experiments were carried out with anti-phosphotyrosine antibodies. Both EGF and TGF- $\alpha$  (34 nM) increased the phosphotyrosine content of HER497K (Fig. 4A). Maximal effects were seen 2–5 min after addition of either EGF or TGF- $\alpha$  (Fig. 4A). To confirm that the phosphorylated band represented HER497K and that HER497R was also phosphorylated, EGFR was immunoprecipitated from CHO cells expressing HER497K and HER497R using an anti-EGFR antibody (13A9), followed by immunoblotting with the anti-phosphotyrosine antibodies.

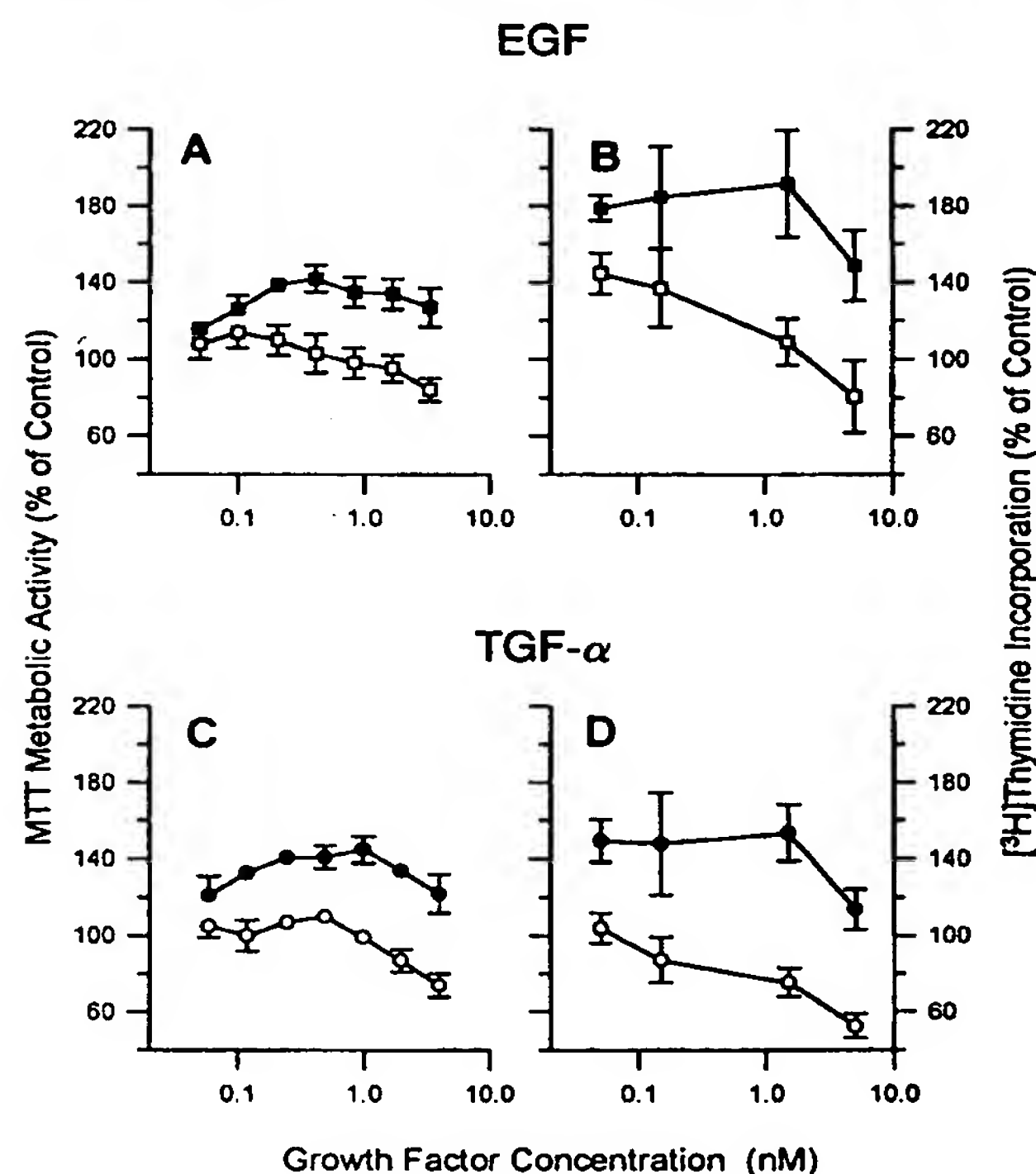
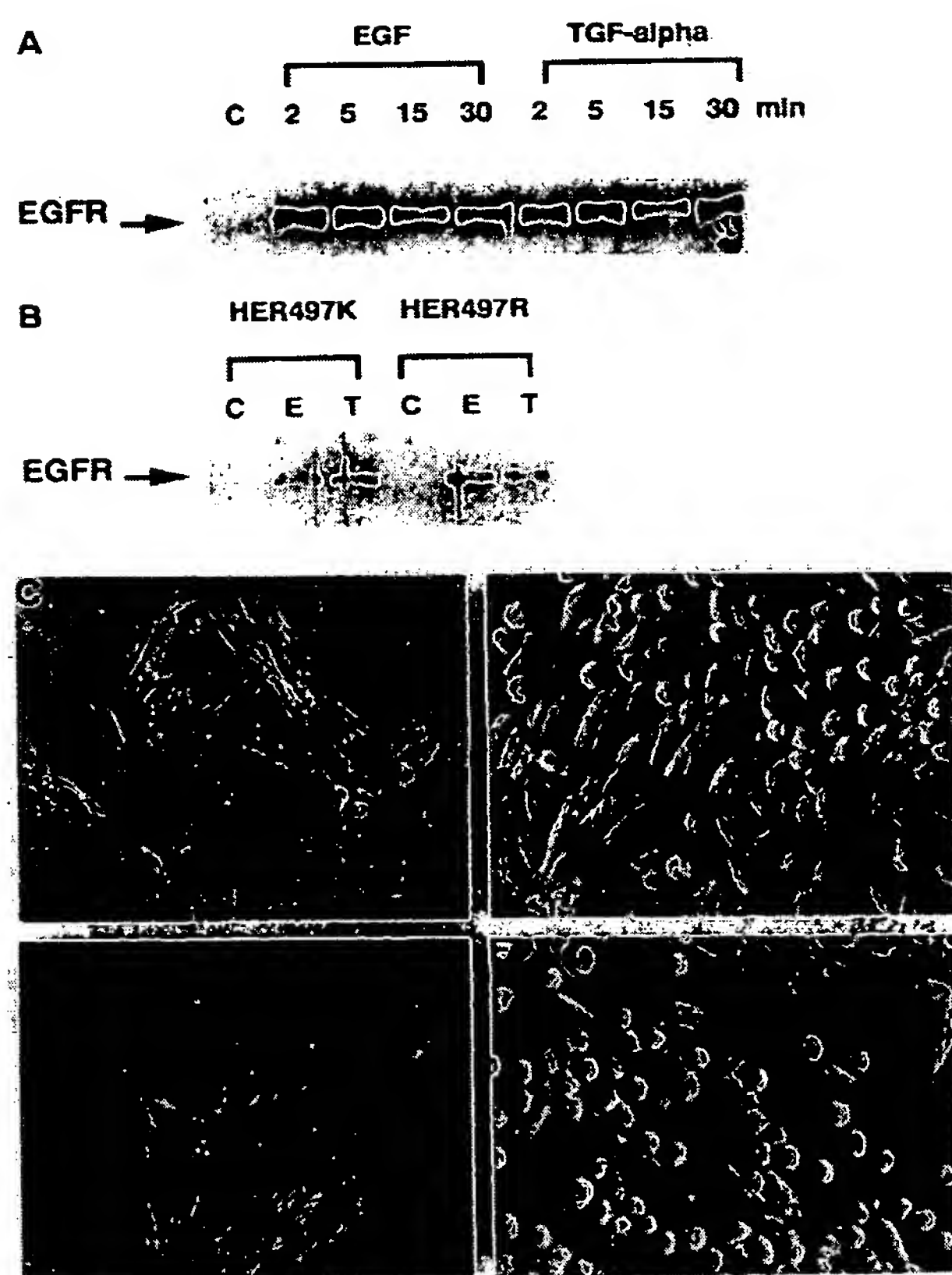


FIG. 3. Effects of EGF and TGF- $\alpha$  on cell growth. HER497R-expressing (solid symbols) and HER497K-expressing (open symbols) cells were incubated for 72 hr in the absence or presence of the indicated concentrations of EGF and TGF- $\alpha$ . Cell proliferation was determined by the MTT dye reduction assay (A and C) and by monitoring [ $^3\text{H}$ ]thymidine incorporation (B and D). Control values, determined in the absence of growth factor, were defined as 0. Data are expressed as percentage change from control values and are means  $\pm$  SE of quintuplicate determinations from three experiments (MTT) or triplicate determinations from three experiments ([ $^3\text{H}$ ]thymidine incorporation).



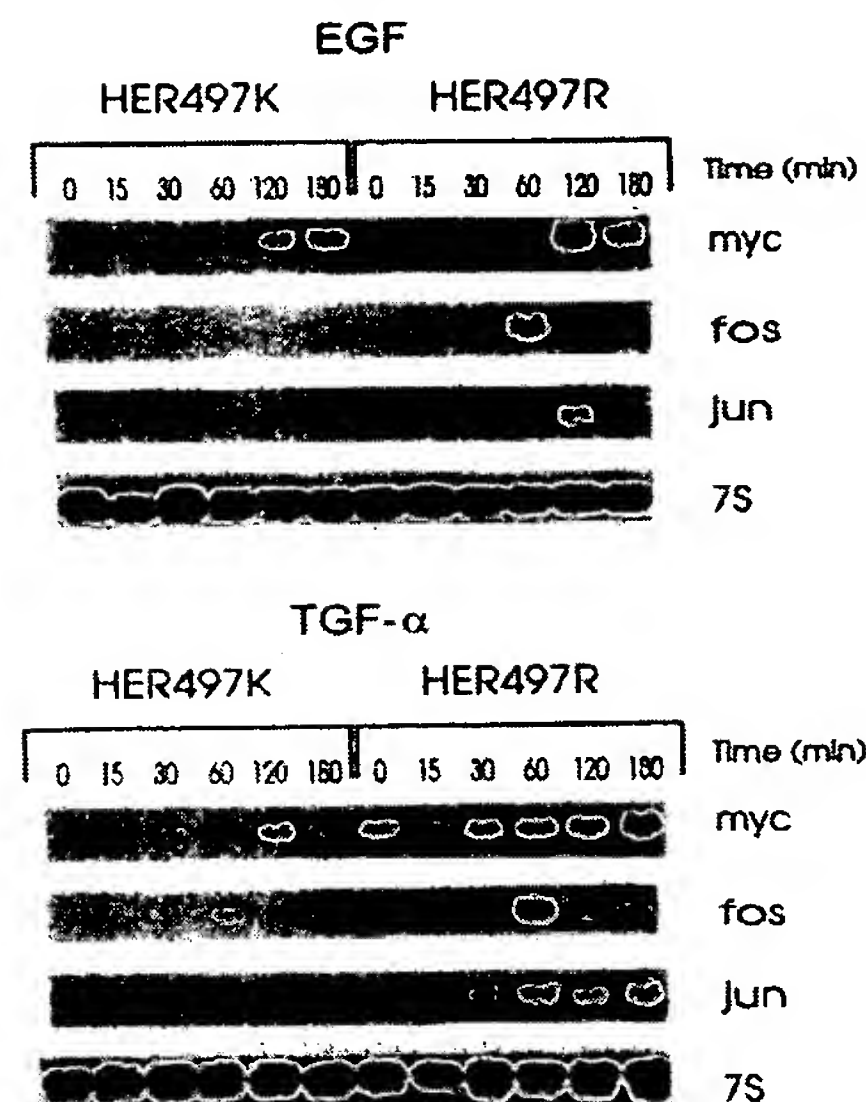


**FIG. 4.** Effects of EGF and TGF- $\alpha$  on phosphotyrosine content and cell morphology. (A) HER497K-expressing cells were incubated at 37°C with 34 nM EGF or TGF- $\alpha$  for the indicated times. Samples were lysed, subjected to SDS/PAGE, and transferred to membranes. The membranes were incubated with the PY-20 anti-phosphotyrosine antibodies and treated with ECL Western blotting reagents. (B) HER497R- and HER497K-expressing cells were incubated for 2 min at 37°C with 3.4 nM EGF (lanes E) or TGF- $\alpha$  (lanes T). HER497R and HER497K were then immunoprecipitated, and the immunoprecipitates were subjected to immunoblotting with PY-20 antibodies as described above. Lanes C, control. (C-F) Cell morphology. HER497R- (C and D) and HER497K- (E and F) expressing cells were grown in six-well plates in medium containing 0.5% FBS in the absence (C and E) or presence (D and F) of 3.4 nM EGF. ( $\times 320$ .)

Both EGF and TGF- $\alpha$  (3.4 nM) enhanced tyrosine phosphorylation of HER497K and HER497R (Fig. 4B), even at concentrations as low as 250 pM (data not shown).

**Effects of EGF and TGF- $\alpha$  on *MYC*, *FOS*, and *JUN* mRNA Levels and Cell Morphology.** In both HER497K- and HER497R-expressing cells, EGF (1.7 nM) caused a time-dependent increase in the mRNA levels of the protooncogenes *MYC*, *FOS*, and *JUN*. However, this effect was consistently more pronounced in the case of HER497R by comparison with HER497K (Fig. 5). Similarly, the effects of TGF- $\alpha$  on the expression of these mRNA moieties was more pronounced in the HER497R-expressing cells (Fig. 5). When two additional HER497R-expressing clones were examined for *MYC*, *FOS*, and *JUN* induction, EGF and TGF- $\alpha$  again exerted a greater effect on the expression of these mRNA moieties by comparison with several HER497K-expressing clones (data not shown).

Parental CHO cells were pleomorphic and flat and did not alter their shape after EGF or TGF- $\alpha$  addition (data not shown). HER497R-expressing cells had a polygonal shape and acquired either a fusiform or a rounded morphology after



**FIG. 5.** Effects of EGF and TGF- $\alpha$  on *MYC*, *FOS*, and *JUN* mRNA levels. HER497R- and HER497K-expressing cells were incubated with 1.7 nM EGF or TGF- $\alpha$  for the indicated times. RNA was isolated and analyzed by Northern blotting. Exposure times: *MYC*, 48 hr; *FOS*, 96 hr; *JUN*, 24 hr; 7S, 16 hr. Each autoradiograph is representative of at least three experiments.

addition of EGF for 40 hr (Fig. 4 C and D). HER497K-expressing cells had a uniformly narrow and fusiform shape in the absence of EGF (Fig. 4E), with many cells acquiring a rounded morphology after incubation with EGF for 40 hr (Fig. 4F). Similar morphological differences were observed with TGF- $\alpha$  as well as in two other HER497R- and HER497K-expressing clones (data not shown).

## DISCUSSION

In the present study, we expressed variant (HER497K) and wild-type (HER497R) EGFRs in CHO cells that are devoid of endogenous EGFRs and compared the two receptors with respect to ligand binding, growth stimulation, tyrosine kinase activation, and induction of *MYC*, *FOS*, and *JUN*. There were two major differences in the binding characteristics of HER497R- and HER497K-expressing CHO cells. First, irrespective of whether binding was performed at pH 7.4 or 8.0, there were no high-affinity binding sites for  $^{125}$ I-TGF- $\alpha$  with HER497K. In contrast, similar high- and low-affinity binding sites were obtained with  $^{125}$ I-EGF and  $^{125}$ I-TGF- $\alpha$  in the case of HER497R and with  $^{125}$ I-EGF in the case of HER497K. Second, at high pH levels, TGF- $\alpha$  dissociation was more pronounced with HER497K by comparison with HER497R. Thus, a single amino acid substitution in subdomain IV of EGFR selectively alters TGF- $\alpha$  binding and dissociation, even though the EGF/TGF- $\alpha$  binding region is located in a 47-amino acid sequence (residues 321–367) within subdomain III (residues 313–446) of EGFR (6, 7). Inasmuch as EGFR dimers appear to have a higher ligand binding affinity than EGFR monomers (17), our findings raise the possibility that the arginine-to-lysine substitution may result in attenuated TGF- $\alpha$ -induced HER497K dimerization. They also imply a modulatory role for this EGFR region on the TGF- $\alpha$  binding pocket within subdomain III.

Tyrosine phosphorylation of specific amino acids in the intracellular domain of EGFR allows for recruitment of important regulatory intracellular substrates such as Shc and Grb2, activation of p21<sup>ras</sup>, initiation of a cascade of phos-



phorylation reactions that activate cytosolic proteins, and activation of nuclear oncoproteins (2, 18). In this study, EGF and TGF- $\alpha$  efficiently increased tyrosine phosphorylation of HER497R and HER497K, indicating that the intrinsic tyrosine kinase activity of both receptors was activated by both ligands. However, in HER497R-expressing cells, EGF and TGF- $\alpha$  significantly increased cell growth and caused the anticipated increases in *MYC*, *FOS*, and *JUN* mRNA levels. In contrast, in HER497K-expressing cells, both factors exhibited a markedly reduced growth response and a decreased ability to raise *MYC*, *FOS*, and *JUN* mRNA levels. These differences indicate that, in spite of relatively similar intrinsic tyrosine kinase autophosphorylating activities, HER497K activates a less mitogenic pathway than HER497R. Therefore, it is possible that the region in the extracellular domain of EGFR that contains the arginine-to-lysine substitution somehow modulates the efficiency of recruitment of intracellular substrates to the activated receptor and/or allows for the activation of alternative signaling pathways that do not lead to efficient nuclear protooncogene induction and subsequent growth stimulation. The hypothesis that this portion of EGFR may represent an important regulatory region is supported by the observations that the seven residues upstream of codon 497 and the five downstream residues are conserved in the mouse, rat, and human EGFRs (19–21) and that the intercysteine spacing is also completely conserved in this region of the human, rat, mouse, and chicken EGFRs as well as in *let-23*, the homologue of EGFR in *Caenorhabditis elegans* (22).

EGF and TGF- $\alpha$  generally exert the same effects on cell proliferation and function in a variety of cell types. However, a number of quantitative and qualitative differences have been reported between their biological actions (9, 23, 24). In the present study, HER497K-expressing cells were markedly growth inhibited by high concentrations of TGF- $\alpha$  but not EGF, perhaps because of the absence of high-affinity TGF- $\alpha$  binding sites in these cells. Neither factor inhibited the growth of HER497R-expressing cells. Therefore, our findings raise the possibility that some of the differences between the actions of EGF and TGF- $\alpha$  may be due to the presence of HER497K or other variant EGFRs in these cells. In support of this hypothesis, TGF- $\alpha$  exerts a more potent growth inhibitory effect than EGF in RL95-2 human endometrial cancer cells (23), which express both HER497R and HER497K (8).

Recently, several null alleles that result in loss of *let-23* function have been identified (22). Two of these alterations are located in subdomain IV and involve mutations of cysteine to either tryptophan or tyrosine at codons 700 and 753, respectively (22). It is unlikely that these null mutants are analogous to the variant human EGFR described in the present study, inasmuch as HER497K kinase activity was intact and cell proliferation and nuclear protooncogene induction were regulated by this variant receptor in response to ligand binding. In addition, both HER497K and HER497R altered the flattened, nontransformed morphology of the

parental CHO cells. Taken together, our data suggest that the extracellular juxtamembranous region of human EGFR has the potential to modulate both ligand binding and transmembrane signaling to the intracellular domain, which may allow the EGF family of ligands to differentially regulate a variety of cellular processes.

1. Czech, M. P. (1989) *Cell* 59, 235–238.
2. Schlessinger, J. & Ullrich, A. (1992) *Neuron* 9, 383–391.
3. Thorne, B. A. & Plowman, G. D. (1994) *Mol. Cell. Biol.* 14, 1635–1646.
4. Shoelson, S. E., White, M. F. & Kahn, C. R. (1988) *J. Biol. Chem.* 263, 4852–4860.
5. Hsu, C. Y., Mohammadi, M., Nathan, M., Honegger, A. M., Ullrich, A., Schlessinger, J. & Hurwitz, D. R. (1990) *Cell Growth Differ.* 1, 191–200.
6. Lax, I., Burgess, W. H., Bellot, F., Ullrich, A., Schlessinger, J. & Givol, D. (1988) *Mol. Cell. Biol.* 8, 1831–1834.
7. Wu, D., Wang, L., Chi, Y., Sato, G. H. & Sato, J. D. (1990) *Proc. Natl. Acad. Sci. USA* 87, 3151–3155.
8. Moriai, T., Kobrin, M. S. & Korc, M. (1993) *Biochem. Biophys. Res. Commun.* 191, 1034–1039.
9. Winkler, M. E., O'Connor, L., Winget, M. & Fendly, B. (1989) *Biochemistry* 28, 6373–6378.
10. Chen, C. & Okayama, H. (1987) *Mol. Cell. Biol.* 7, 2745–2752.
11. Korc, M. & Finman, J. E. (1989) *J. Biol. Chem.* 264, 14990–14999.
12. Munson, P. J. & Robard, D. (1980) *Anal. Biochem.* 107, 220–239.
13. Raitano, A. B., Scuderi, P. & Korc, M. (1990) *Pancreas* 5, 267–277.
14. Korc, M., Chandrasekar, B. & Shah, G. N. (1991) *Cancer Res.* 51, 6243–6249.
15. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. T. & Rutter, W. J. (1979) *Biochemistry* 18, 5294–5299.
16. Korc, M., Chandrasekar, B., Yamanaka, Y., Friess, H., Buchler, M. & Beger, H. G. (1992) *J. Clin. Invest.* 90, 1352–1360.
17. Yarden, Y. & Schlessinger, J. (1987) *Biochemistry* 26, 1443–1451.
18. Gotoh, N., Tojo, A., Muroya, K., Hashimoto, Y., Hattori, S., Nakamura, S., Takenawa, T., Yazaki, Y. & Shibuya, M. (1994) *Proc. Natl. Acad. Sci. USA* 91, 167–171.
19. Avivi, A., Lax, I. M., Ullrich, A., Schlessinger, J., Givol, D. & Morse, B. (1991) *Oncogene* 6, 673–676.
20. Lax, I., Johnson, A., Howk, R., Sap, J., Bellot, F., Winkler, M., Ullrich, A., Vennstrom, B., Schlessinger, J. & Givol, D. (1988) *Mol. Cell. Biol.* 8, 1970–1978.
21. Petch, L. A., Harris, J., Raymond, V. W., Blasband, A., Lee, D. C. & Earp, H. S. (1990) *Mol. Cell. Biol.* 10, 2973–2982.
22. Aroian, R. V., Lessa, G. M. & Sternberg, P. W. (1994) *EMBO J.* 13, 360–366.
23. Korc, M., Haussler, C. A. & Trookman, N. S. (1987) *Cancer Res.* 47, 4909–4914.
24. Gan, B. S., Hollenberg, M. D., MacCannell, K. L., Lederis, K., Winkler, M. E. & Derynck, R. (1987) *J. Pharmacol. Exp. Ther.* 242, 331–337.
25. Ullrich, A., Coussens, L., Hayflick, J. S., Dull, T. J., Gray, A., Tam, A. W., Lee, J., Yarden, Y., Libermann, T. A., Schlessinger, J., Downward, J., Mayes, E. L. V., Whittle, N., Waterfield, M. D. & Seeburg, P. H. (1984) *Nature (London)* 309, 418–425.

# INHIBITION OF CYTOTOXIC T CELL DEVELOPMENT BY TRANSFORMING GROWTH FACTOR $\beta$ AND REVERSAL BY RECOMBINANT TUMOR NECROSIS FACTOR $\alpha$

By GERALD E. RANGES, IRENE S. FIGARI, TERJE ESPEVIK, AND  
MICHAEL A. PALLADINO, JR.

*From the Department of Molecular Immunology, Genentech, Inc.,  
South San Francisco, California 94080*

Transforming growth factor  $\beta$  (TGF- $\beta$ ),<sup>1</sup> although first defined for its ability to induce nonneoplastic cells to express a transformed phenotype, has now been shown to exert multiple actions on both normal and transformed cells (1). Recent studies have also defined a variety of immunoregulatory properties of TGF- $\beta$ , including inhibition of T and B cell proliferation, IL-2-R induction, IL-1-induced thymocyte proliferation, cytokine production, including IFN- $\gamma$  and TNF- $\alpha$ , natural killer cell activity, and class II antigen expression (2-6). The mechanism(s) through which TGF- $\beta$  exerts these immunoregulatory effects is at present not known.

TNF- $\alpha$  (also referred to as cachectin) has also been shown to express multifunctional immunomodulatory activities besides its direct cytotoxic/cytostatic effects on transformed cells (7-10). However, in contrast, to TGF- $\beta$ , TNF- $\alpha$  enhances IL-2-R expression, class II antigen expression, and IFN- $\gamma$  production by activated lymphocytes (9, 10).

The contrasting immunoregulatory activities of these two proteins prompted studies to further define their immunoregulatory activities in vitro. In this report, we describe the dose-dependent inhibition of CTL generation by TGF- $\beta$  and the reversal of this inhibition by recombinant murine TNF- $\alpha$  (rMuTNF- $\alpha$ ). In addition, we demonstrate that TNF- $\alpha$  is an important cytokine involved in CTL development.

## Materials and Methods

**Animals.** 6-12-wk-old female BALB/c and C57BL/6 (B6) mice were obtained from Charles River Breeding Laboratories (Wilmington, MA).

**Reagents.** Porcine platelet-derived TGF- $\beta$  (R and D Systems, Minneapolis, MN) was reconstituted in 4 mM HCl to 1  $\mu$ g/ml and stored at 4°C. rMuTNF- $\alpha$  (sp act  $7 \times 10^7$  U/mg), as determined by a standard cytotoxic bioassay using L-M cells, contained  $<0.025$  pg of endotoxin per microgram protein by the limulus amoebocyte assay (11, 12). The specific activities of recombinant human TNF- $\alpha$  (rHuTNF- $\alpha$ ) and - $\beta$  provided by Genentech, Inc., as determined by the L-M bioassay were  $5 \times 10^7$  and  $2 \times 10^8$  U/mg protein, respectively (13, 14).

<sup>1</sup> *Abbreviations used in this paper:* CMEM, complete minimal essential medium; NRS, normal rabbit serum; TGF- $\beta$ , transforming growth factor  $\beta$ .

T. Espevik is a visiting scientist from the Institute of Cancer Research, University of Trondheim, Norway.

**Mixed Lymphocyte Cultures (MLC).** CTLs were generated in 5-d MLC by incubation in 24-well tissue culture plates (3524; Costar, Cambridge, MA) of  $5 \times 10^6$  or  $5 \times 10^5$  B6 responding spleen cells and  $5 \times 10^6$  BALB/c stimulator spleen cells (irradiated 2,000 rad) per well in 2 ml of complete minimal essential medium (CMEM) consisting of Eagle's minimum essential medium supplemented with 0.1 mM nonessential amino acids, 2 mM L-glutamine, penicillin (100 U/ml), and streptomycin (100  $\mu$ g/ml) (Gibco, Grand Island, NY) and 10% heat-inactivated FCS (Hyclone Laboratories, Inc., Logan, UT).

**CTL Assay.** After 5 d of culture, the cells were harvested, washed three times in CMEM and tested for cytotoxic activity in a 4-h  $^{51}\text{Cr}$ -release assay. P815 (DBA/2 mastocytoma, H-2<sup>d</sup>; American Type Culture Collection, Rockville, MD) or LBRM-33-1A5B (B10.BR lymphoma, H-2<sup>k</sup>; A. Zlotnik, DNAX Research Institute for Molecular and Cellular Biology, Palo Alto, CA) were labeled with 150  $\mu\text{Ci}$   $\text{Na}^{51}\text{CrO}_4$  (5 mCi/ml; Amersham Corp., Arlington Heights, IL) for 45 min at 37°C, followed by three washes in CMEM. 100  $\mu\text{l}$  of target cells ( $10^5$  cells/ml) and 100  $\mu\text{l}$  of effector cells at various concentrations were added in triplicate in 96-well round-bottom microtiter plates (Costar). After 4 h of incubation at 37°C and 5%  $\text{CO}_2$  in air, the supernatants were harvested (Skatron, Rockville, MD) and their radioactivity was determined in an automatic gamma counter (Micromedic Systems, Horsham, PA). Percent specific cytotoxicity was calculated as  $100 \times [\text{cpm of test supernatants of effector cells and target cells incubated together (experimental release)}] - [\text{cpm of supernatants of target cells incubated alone (spontaneous release)}] / [\text{cpm after lysis of target cells with 2\% NP-40 (maximum release)}] - [\text{spontaneous release}]$ . Results given are the mean of triplicate cultures  $\pm$  SE. Spontaneous release of target cells alone was  $<10\%$  of maximum for all experiments.

**Detection of MuTNF- $\alpha$  in Supernatants of the MLC.** The quantity of MuTNF- $\alpha$  produced in the MLC was determined using the WEHI-164 clone 13, which is capable of detecting levels as low as  $7 \times 10^{-4}$  U/ml (0.01 pg/ml) of MuTNF- $\alpha$  per milliliter (15). Using WEHI-164 clone 13 as indicator cells, the cytotoxicity induced by MuTNF- $\alpha$  was determined by an MTT tetrazolium colorimetric assay as detailed previously (15). The amount of MuTNF- $\alpha$  in test supernatants was calculated on the basis of cytotoxicity obtained in the presence of various dilutions of a rMuTNF- $\alpha$  standard. Data are expressed as units of TNF- $\alpha$  per milliliter.

The specificity of this assay for MuTNF- $\alpha$  was demonstrated by inhibiting the cytolytic activity of test supernatants with anti-rMuTNF- $\alpha$  antiserum produced in New Zealand rabbits. This antiserum developed by methods similar to those described for production of anti-rHuTNF- $\alpha$  antiserum (16), had a neutralization titer of  $\sim 10^6$  U/ml in the L-M bioassay (12). Normal rabbit serum (NRS) was obtained from control New Zealand rabbits.

## Results

**Inhibition of CTL Generation by TGF- $\beta$ .** When added to MLC, TGF- $\beta$  inhibited the generation of B6-anti-H-2<sup>d</sup>-specific CTL in a dose-dependent fashion (Table I). The inhibitory effects were most pronounced when TGF- $\beta$  was added during the first 24 h of the MLC (Fig. 1). Less inhibition was observed when the addition was delayed for 48 h, and only a minimal effect was seen if the delay was more than 72 h. Similarly, at the doses tested TGF- $\beta$  showed no inhibitory activity if included directly in the CTL assay. These data indicate that TGF- $\beta$  inhibits CTL generation in a dose-dependent manner, and its mechanism of action appears to involve early stages of the MLC.

**Inhibition of TNF- $\alpha$  Production by TGF- $\beta$  during MLC.** Although a variety of in vitro immunoregulatory activities have been ascribed to TNF- $\alpha$ , studies examining the production of TNF- $\alpha$  during an MLC and the effects of altering endogenous TNF- $\alpha$  levels during CTL development have not been reported (8–10). Moreover, our previous studies have shown that TGF- $\beta$  can inhibit TNF- $\alpha$  production by murine macrophages (6). Therefore, we considered it important



TABLE I  
Effect of TGF- $\beta$  on In Vitro CTL Generation

TGF- $\beta$ * (ng/ml)	Percent specific $^{51}\text{Cr}$ release ( $\pm$ SE) at <sup>‡</sup> E/T ratios of:			
	50:1	25:1	12.5:1	6.25:1
None	89 $\pm$ 2.0	89 $\pm$ 4.0	86 $\pm$ 0.4	75 $\pm$ 1.0
10	18 $\pm$ 2.0	8 $\pm$ 1.0	7 $\pm$ 2.0	2 $\pm$ 0.5
1.0	30 $\pm$ 0.4	16 $\pm$ 2.0	8 $\pm$ 1.0	3 $\pm$ 1.0
0.10	82 $\pm$ 0.4	79 $\pm$ 2.0	82 $\pm$ 0.2	63 $\pm$ 1.0

\* TGF- $\beta$  was added at the start of culture.

<sup>‡</sup> P815 (H-2<sup>d</sup>) target cells were used. Data are from one of three representative experiments. Percent lysis against  $^{51}\text{Cr}$ -labeled LBRM-33-1A5B was <10% at 50:1 E/T ratio.

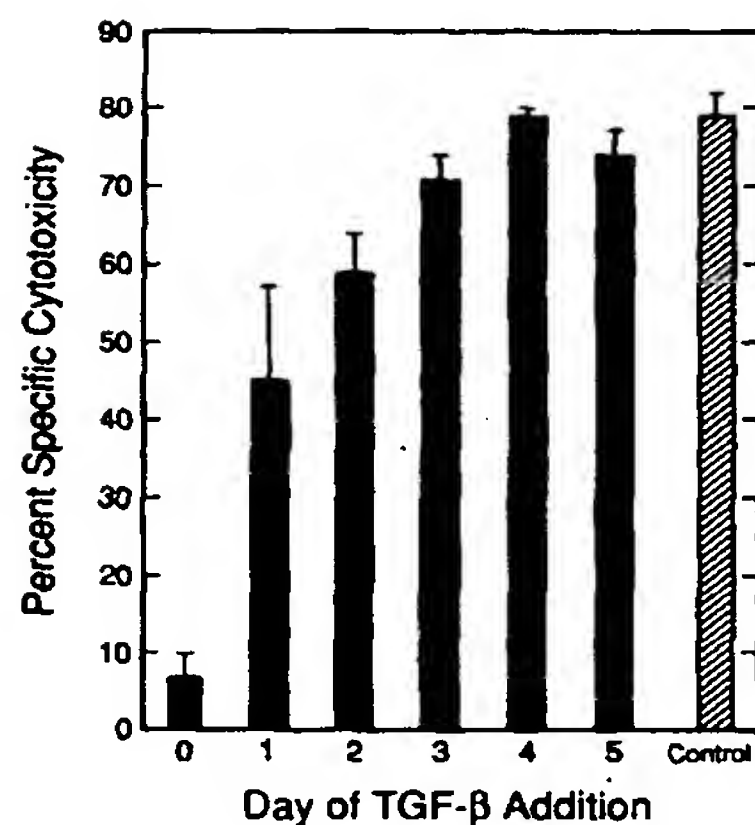


FIGURE 1. Kinetics of TGF- $\beta$  inhibition of CTL generation. 10 ng/ml TGF- $\beta$  were added to MLC on days shown. On day 5, this same concentration of TGF- $\beta$  was added to the CTL assay. Data presented are the mean  $\pm$  SE of four independent experiments performed at an E/T ratio of 25:1. Similar results were obtained at all E/T ratios tested (data not shown).

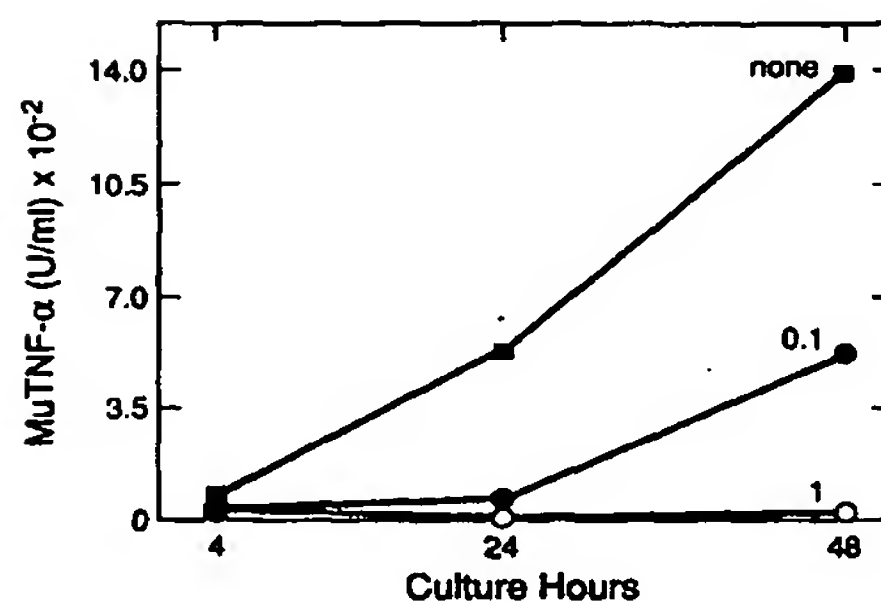


FIGURE 2. Inhibition of TNF- $\alpha$  production during MLC by TGF- $\beta$ . 300  $\mu$ l were removed from replicate MLC at the times indicated, and assayed for TNF- $\alpha$  as described in Materials and Methods. MLC contained no TGF- $\beta$  (■), 0.1 ng TGF- $\beta$ /ml (●), or 1.0 ng TGF- $\beta$ /ml (○). The presence of 10  $\mu$ g TGF- $\beta$  in the WEHI-164 MuTNF- $\alpha$  bioassay did not affect the measurement of MuTNF- $\alpha$  activity, while the presence of anti-rMuTNF- $\alpha$  antiserum reduced the detection of MuTNF- $\alpha$  to background levels. Standard errors for all determinations were <10%. Results are presented as mean of triplicate determinations of one of four representative experiments. No significant amounts of MuTNF- $\alpha$  were produced by cultures containing B6 cells or irradiated BALB/c cells alone.

to examine whether the inhibitory effects of TGF- $\beta$  on CTL development could in part be due to the inhibition of TNF- $\alpha$  production. The concentration of MuTNF- $\alpha$  was measured at various times during the first 48 h of MLC in the absence or presence of 1.0 or 0.1 ng/ml TGF- $\beta$ . The data (Fig. 2) indicate that as early as 4 h after culture initiation,  $\sim 7 \times 10^{-5}$  U/ml (0.1 pg/ml) of TNF- $\alpha$

TABLE II  
Reversal of TGF- $\beta$  Inhibition of CTL Generation by TNF- $\alpha$

Treatment*	Percent specific $^{51}\text{Cr}$ release ( $\pm$ SE) at <sup>‡</sup> E/T ratios of:		
	25:1	12.5:1	6.25:1
Control	83 $\pm$ 5.0	80 $\pm$ 4.0	69 $\pm$ 4.0
10 ng/ml TGF- $\beta$	16 $\pm$ 1.0	12 $\pm$ 3.0	7 $\pm$ 2.0
10 <sup>4</sup> U/ml rMuTNF- $\alpha$	83 $\pm$ 2.0	85 $\pm$ 4.0	78 $\pm$ 1.0
10 ng/ml TGF- $\beta$ + rMuTNF- $\alpha$ at			
10 <sup>5</sup> U/ml	54 $\pm$ 7.0	35 $\pm$ 1.0	21 $\pm$ 0.6
10 <sup>4</sup> U/ml	62 $\pm$ 3.0	48 $\pm$ 1.0	25 $\pm$ 2.0
10 <sup>3</sup> U/ml	49 $\pm$ 3.0	40 $\pm$ 3.0	18 $\pm$ 1.0
10 <sup>2</sup> U/ml	28 $\pm$ 1.0	20 $\pm$ 3.0	11 $\pm$ 2.0
10 <sup>1</sup> U/ml	16 $\pm$ 2.0	8 $\pm$ 0.4	5 $\pm$ 1.0

\* TGF- $\beta$  and/or rMuTNF- $\alpha$  were added to MLC on day 0 at concentrations indicated.

<sup>‡</sup> Data are from one of three representative experiments.

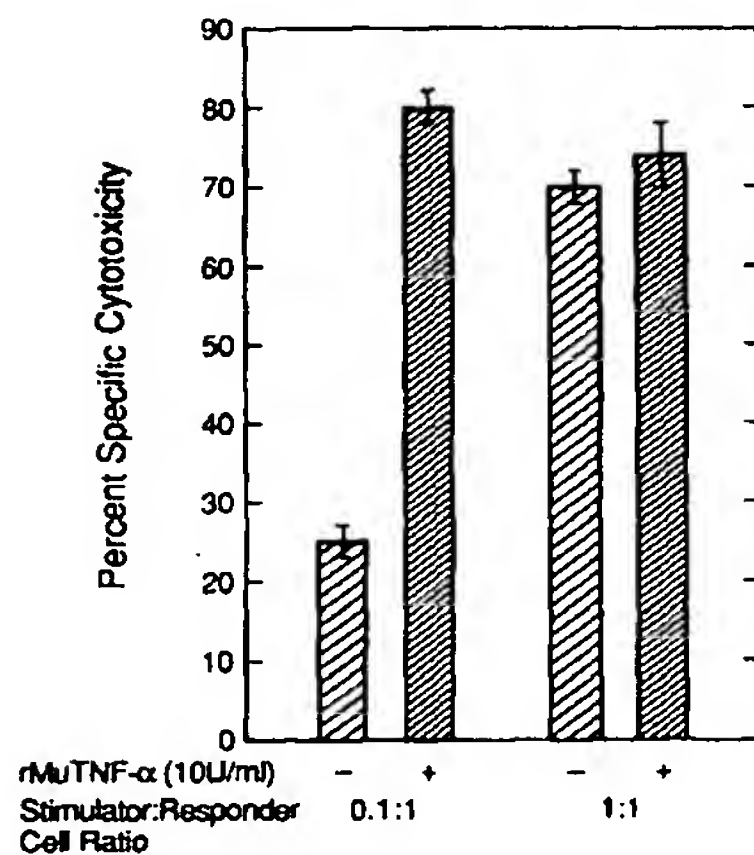


FIGURE 3. Effects of rMuTNF- $\alpha$  on CTL generation during suboptimal culture conditions. MLC were established at stimulator/responder ratios of 1:1 ( $\square$ ) and 0.1:1 ( $\square$ ). 10 U/ml rMuTNF- $\alpha$  was added on day 0 to MLC at a 0.1:1 ratio ( $\square$ ) and at a 1:1 ratio ( $\blacksquare$ ). Results are mean  $\pm$  SE of triplicate determinations performed at an E/T cell ratio of 12.5:1. Similar results were obtained at all E/T cell ratios tested (data not shown).

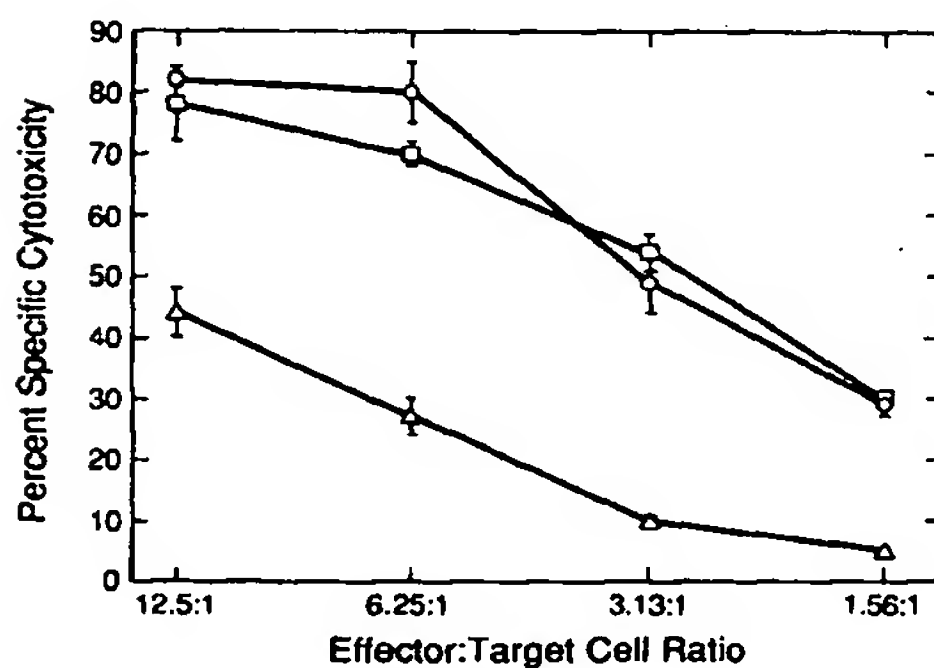


FIGURE 4. Inhibition of CTL generation by antibodies to rMuTNF- $\alpha$ . Rabbit serum polyclonal antibodies to rMuTNF- $\alpha$  was added at a 1:100 final dilution on day 0 of MLC. MLC contained: no antibodies ( $\square$ ); polyclonal antibodies to rMuTNF- $\alpha$  ( $\Delta$ ); NRS ( $\circ$ ). Results are mean  $\pm$  SE of triplicate determinations. CTL were washed three times in CMEM to prevent carryover of antibodies to rMuTNF- $\alpha$  into the  $^{51}\text{Cr}$  assay.

can be detected and that TGF- $\beta$  suppresses the production of TNF- $\alpha$  in a dose-dependent manner.

To determine if the suppression of MuTNF- $\alpha$  production was critical to the inhibitory activity of TGF- $\beta$  on CTL generation, exogenous rMuTNF- $\alpha$  was added to MLC in the absence and presence of TGF- $\beta$  (Table II). The addition of  $10^2$  U/ml of rMuTNF- $\alpha$  to TGF- $\beta$ -suppressed MLC significantly restored the CTL activity. These results are not due to any direct activity of rMuTNF- $\alpha$  on the P815 target cells since the viability of these cells was not affected by rMuTNF- $\alpha$  doses as high as  $10^5$  U/ml. When rHuTNF- $\alpha$  was substituted for rMuTNF- $\alpha$ , reversal of TGF- $\beta$  suppression was detected only at concentrations of  $10^5$  U/ml rHuTNF- $\alpha$ , and no activity was observed when rHuTNF- $\beta$  was used at this same concentration (data not shown). These results may indicate a species preference with regard to action of TNF- $\alpha$  in this context, as suggested in previous studies (11).

*Enhancement of CTL Development by rMuTNF- $\alpha$ .* Under the optimal conditions used in developing the CTLs, the addition of rMuTNF- $\alpha$  at the doses tested had only minimal enhancing activity. However, at a suboptimal responder/stimulator ratio (0.1:1), rMuTNF- $\alpha$  significantly enhanced both the proliferative response (data not shown) and CTL generation to H-2<sup>d</sup> targets (Fig. 3).

*Inhibition of CTL Generation by Antibodies to rMuTNF- $\alpha$ .* As our earlier studies demonstrated that rMuTNF- $\alpha$  can significantly enhance CTL development if added to MLC established at suboptimal stimulator to responder ratios, we considered it important to investigate whether antibodies to rMuTNF- $\alpha$  could inhibit CTL generation. As shown in Fig. 4, the addition of rabbit polyclonal antibodies to rMuTNF- $\alpha$ , but not addition of NRS to MLC on day 0 significantly inhibited CTL generation, further supporting our data of the critical role of TNF- $\alpha$  during CTL development.

### Discussion

We have investigated the effects of TGF- $\beta$  and rMuTNF- $\alpha$  on CTL generation and function. Our investigations have indicated that: (a) TGF- $\beta$  will inhibit, in a dose-dependent manner, CTL generation, but only when TGF- $\beta$  is added in the early stages of the MLC; (b) TGF- $\beta$  does not inhibit the cytotoxic activity of CTL; (c) TNF- $\alpha$  production, which can be detected as early as 4 h after the initiation of an MLC, is inhibited by TGF- $\beta$ ; (d) addition of rMuTNF- $\alpha$  to TGF- $\beta$ -inhibited MLC significantly reverses the inhibitory activity of TGF- $\beta$ ; and (e) antibodies to rMuTNF- $\alpha$  significantly inhibit CTL development. These results show that TGF- $\beta$  does not inhibit CTL generation by nonspecific cytostatic or cytotoxic processes, since the addition of TGF- $\beta$  to MLC for the final 72 h does not affect the CTL activity. Rather, TGF- $\beta$  appears to inhibit a differentiation step in the early development of CTL. In addition, similar results have been obtained with rHuTGF- $\beta$  (the polypeptide sequence of mature human and porcine TGF- $\beta$  is identical) indicating that the effects observed with the natural porcine preparation are mediated by TGF- $\beta$  alone, and not a contaminant (data not shown) (17).

Previous studies have indicated that CTL generation involves the production



of two or more lymphokines, including IL-2 (18-21). TGF- $\beta$  has been shown to affect T cell activation by downregulating IL-2-R expression (2) and thereby inhibiting T cell proliferation. Scheurich et al. (10) have reported that rHuTNF- $\alpha$  can enhance the response to IL-2 by upregulating the expression of IL-2-R. Our data indicate that TNF- $\alpha$  plays an important role in the generation of CTL, and that TGF- $\beta$  may inhibit the development of such cells, at least in part, by inhibiting TNF- $\alpha$  production. Our data also demonstrate that, while the reversal of TGF- $\beta$  suppression is significant, it is not complete, regardless of the rMuTNF- $\alpha$  dose used. This may point to multiple effects of TGF- $\beta$  on CTL generation, only some of which are reversible by TNF- $\alpha$ , and which require the presence of additional cytokines. In recent experiments, we have observed that IL-2, which induces TNF- $\alpha$  production, can also reverse TGF- $\beta$  inhibition of CTL generation (our unpublished observations) (22). Thus, TGF- $\beta$  may inhibit CTL generation directly by downregulating IL-2-R expression, or indirectly by preventing up-regulation of this receptor by blocking TNF- $\alpha$  production. Studies are now under way to further investigate the role of TNF- $\alpha$  in CTL generation, and the suppression of that process by TGF- $\beta$ .

### Summary

The immunoregulatory effects of transforming growth factor  $\beta$  (TGF- $\beta$ ) and recombinant murine tumor necrosis factor  $\alpha$  (rMuTNF- $\alpha$ ) on CTL generation and activity were examined. The results demonstrate that TGF- $\beta$ , in a dose-dependent manner, inhibited CTL generation but not CTL activity. The inhibitory effects were detected only when TGF- $\beta$  was added within the first 48 h of the MLC. Little activity was seen when it was added thereafter, including the addition of TGF- $\beta$  to the cytotoxicity assay. The production of TNF- $\alpha$ , which occurs during early phases of the MLC and which is inhibited in the presence of TGF- $\beta$ , appears to have an important regulatory role, as altering the levels of TNF- $\alpha$  in an MLC can significantly influence CTL development. The inhibitory effects of TGF- $\beta$  on the MLC can be significantly reversed by the addition of rMuTNF- $\alpha$  to the cultures. These results demonstrate that TGF- $\beta$  can inhibit MLC and subsequent CTL generation at early stages of the reaction, and such inhibition may involve the suppression of TNF- $\alpha$  production.

We thank Ms. Stephanie Shipley for excellent technical assistance, Mr. Chris Nelson for preparing the rabbit anti-rMuTNF- $\alpha$  antisera, and Dr. E. Rinderknecht (Genentech, Inc.) for supplying rMuTNF- $\alpha$ . The authors also thank Ms. Socorro Cuisia for excellent secretarial assistance.

*Received for publication 8 June 1987.*

### References

1. Sporn, M. B., A. B. Roberts, L. M. Wakefield, and R. K. Assoian. 1986. Transforming growth factor- $\beta$ : Biological function and chemical structure. *Science (Wash. DC)*. 233:532.
2. Kehrl, J. H., L. M. Wakefield, A. B. Roberts, S. Jakowlew, M. Alvarez-Mon, R. Derynck, M. B. Sporn, and A. S. Fauci. 1986. Production of transforming growth

- factor  $\beta$  by human T lymphocytes and its potential role in the regulation of T cell growth. *J. Exp. Med.* 163:1037.
3. Ristow, H. J. 1986. BSC-1 growth inhibitor 1 type B transforming growth factor is a strong inhibitor of thymocyte proliferation. *Proc. Natl. Acad. Sci. USA.* 83:5531.
  4. Kehrl, J. H., A. B. Roberts, L. M. Wakefield, S. Jakowlew, M. B. Sporn, and A. S. Fauci. Transforming growth factor- $\beta$  is an important immunomodulatory protein for human B lymphocytes. *J. Immunol.* 137:3855.
  5. Palladino, M. A., C. W. Czarniecki, H. H. Chiu, S. M. McCabe, I. S. Figari, and A. J. Ammann. 1986. Regulation of cytokine production and class II antigen expression by transforming growth factor-beta. *UCLA Symp. Growth Factors.* In press.
  6. Espevik, T., I. S. Figari, M. R. Shalaby, G. A. Lackides, G. D. Lewis, H. M. Shepard, and M. A. Palladino, Jr. 1987. Inhibition of cytokine production by cyclosporin A and transforming growth factor beta. *J. Exp. Med.* 166:571.
  7. Sugarman, B. J., B. B. Aggarwal, P. E. Haas, I. S. Figari, M. A. Palladino, and H. M. Shepard. 1985. Recombinant human tumor necrosis factor-alpha: effects on proliferation of normal and transformed cells in vitro. *Science (Wash. DC).* 230:943.
  8. Abbott, J., P. J. Doyle, K. Ngiam, and C. L. Olson. 1981. Ontogeny of murine lymphocytes I. Maturation of thymocytes induced in vitro by tumor necrosis factor-positive serum (TNF). *Cell Immunol.* 57:237.
  9. Ramila, P., and L. B. Epstein. 1986. Tumor necrosis factor as immunomodulator and mediator of monocyte cytotoxicity induced by itself,  $\gamma$ -interferon and interleukin-1. *Nature (Lond.).* 323:86.
  10. Scheurich, P., B. Thoma, U. Ulcer, and K. Pfizenmaier. 1987. Immunoregulatory activity of recombinant human tumor necrosis factor (TNF) - $\alpha$ : induction of TNF receptors on human T cells and TNF- $\alpha$  mediated enhancement of T cell responses. *J. Immunol.* 138:1786.
  11. Pennica, D., J. S. Hayflick, T. Bringman, M. A. Palladino, and D. V. Goeddel. 1985. Cloning and expression in *E. coli* of the cDNA for murine tumor necrosis factor. *Proc. Natl. Acad. Sci. USA.* 82:6060.
  12. Kramer, S. M., and M. E. Carver. 1986. Serum-free in vitro bioassay for the detection of tumor necrosis factor. *J. Immunol. Methods.* 93:210.
  13. Pennica, D., G. E. Nedwin, J. S. Hayflick, P. H. Seeburg, R. Derynck, M. A. Palladino, W. J. Kohr, B. B. Aggarwal, D. V. Goeddel. 1984. Human tumor necrosis factor: cDNA cloning, expression and homology to lymphotoxin. *Nature (Lond.).* 312:724.
  14. Gray, P. W., B. B. Aggarwal, C. V. Benton, T. S. Bringman, W. J. Henzel, J. A. Jarrett, D. W. Leung, B. Moffat, P. Ng, L. P. Svedersky, M. A. Palladino, and G. E. Nedwin. 1984. Cloning and expression of cDNA for human lymphotoxin, a lymphokine with tumor necrosis activity. *Nature (Lond.).* 312:721.
  15. Espevik, T., and J. Nissen-Meyer. 1986. A highly sensitive cell line, WEHI-164, clone 13, for measuring cytotoxic factor/tumor necrosis factor from human monocytes. *J. Immunol. Methods.* 95:99.
  16. Peters, P. M., J. R. Ortaldo, M. R. Shalaby, L. P. Svedersky, G. E. Nedwin, T. S. Bringman, P. E. Hass, B. B. Aggarwal, R. B. Herberman, D. V. Goeddel, and M. A. Palladino, Jr. 1986. Natural killer-sensitive targets stimulate production of TNF- $\alpha$  but not TNF- $\beta$  (lymphotoxin) by highly purified human large granular lymphocytes. *J. Immunol.* 137:2592.
  17. Derynck, R., and L. Rhee. 1987. Sequence of the porcine transforming growth factor-beta precursor. *Nucleic Acids Res.* 15:3187.
  18. Mannel, D. N., W. Falk, and W. Droge. 1983. Induction of cytotoxic T cell function requires sequential action of three different lymphokines. *J. Immunol.* 130:2508.

998 IMMUNOMODULATION BY TRANSFORMING GROWTH FACTOR  $\beta$

19. Gately, M. K., D. E. Wilson, and H. L. Wong. 1986. Synergy between recombinant interleukin-2 (rIL-2) and IL-2 depleted lymphokine-containing supernatants in facilitating allogeneic human cytolytic T lymphocyte responses in vitro. *J. Immunol.* 136:1274.
20. Yang, S. S., T. R. Malek, M. E. Hargrove, and C. C. Ting. 1985. Lymphokine induced cytotoxicity: Requirement of two lymphokines for the induction of optimal cytotoxic responses. *J. Immunol.* 134:3912.
21. Kanagawa, O., and J. M. Chiller. 1985. Lymphokine-mediated induction of cytotoxic activity in a T cell hybridoma. *J. Immunol.* 134:397.
22. Nedwin, G. E., L. P. Svedersky, T. S. Bringman, M. A. Palladino, Jr., and D. V. Goeddel. 1985. Effect of interleukin-2, interferon- $\gamma$ , and mitogens on the production of tumor necrosis factors  $\alpha$  and  $\beta$ . *J. Immunol.* 135:2492.



**P00761 (TRYP\_PIG)** ★ Reviewed, UniProtKB/Swiss-Prot

Last modified April 5, 2011. Version 100.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Trypsin</b> EC=3.4.21.4
Organism	<b>Sus scrofa (Pig)</b>
Taxonomic identifier	9823 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Laurasiatheria › Cetartiodactyla › Suina › Suidae › Sus

**Protein attributes**

Sequence length	231 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Catalytic activity	Preferential cleavage: Arg- -Xaa, Lys- -Xaa.
Cofactor	Binds 1 calcium ion per subunit.
Subcellular location	Secreted › extracellular space.
Sequence similarities	Belongs to the peptidase S1 family. Contains 1 peptidase S1 domain.

**Ontologies**

<b>Keywords</b>	
Biological process	Digestion
Cellular component	Secreted
Ligand	Calcium Metal-binding
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Zymogen
Technical term	3D-structure Direct protein sequencing

**Gene Ontology (GO)**

Biological process	digestion Inferred from electronic annotation. Source: UniProtKB-KW  proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	extracellular space Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	metal ion binding Inferred from electronic annotation. Source: UniProtKB-KW  serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro
Complete GO annotation...	

**Sequence annotation (Features)**

	Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
<b>Molecule processing</b>						
<input type="checkbox"/>	Propeptide	1 – 8	8	Activation peptide		PRO_0000028205
<input checked="" type="checkbox"/>	Chain	9 – 231	223	Trypsin		PRO_0000028206
<b>Regions</b>						
<input checked="" type="checkbox"/>	Domain	9 – 229	221	Peptidase S1		
<b>Sites</b>						
<input type="checkbox"/>	Active site	48	1	Charge relay system		
<input type="checkbox"/>	Active site	92	1	Charge relay system		
<input type="checkbox"/>	Active site	185	1	Charge relay system		
<input type="checkbox"/>	Metal binding	60	1	Calcium		
<input type="checkbox"/>	Metal binding	62	1	Calcium; via carbonyl oxygen		
<input type="checkbox"/>	Metal binding	65	1	Calcium; via carbonyl oxygen		

<input type="checkbox"/>	Metal binding	70	1	Calcium		
<input type="checkbox"/>	Site	179	1	Required for specificity		
				By similarity		

Amino acid modifications

<input type="checkbox"/>	Disulfide bond	15 ↔ 145				
<input type="checkbox"/>	Disulfide bond	33 ↔ 49				
<input type="checkbox"/>	Disulfide bond	117 ↔ 218				
<input type="checkbox"/>	Disulfide bond	124 ↔ 191				
<input type="checkbox"/>	Disulfide bond	156 ↔ 170				
<input type="checkbox"/>	Disulfide bond	181 ↔ 205				

Natural variations

<input type="checkbox"/>	Natural variant	20	1	I → V		
--------------------------	-----------------	----	---	-------	--	--

Secondary structure

1 ..... 231

■ Helix ■ Strand ■ Turn

Details...

Sequences

Sequence	Length	Mass (Da)
<input type="checkbox"/> P00761 [UniParc]. Last modified July 21, 1986. Version 1. Checksum: A0A125CF7FC138C2	FASTA	231 24,409
<div><div>102030405060</div><div>FPTDDDDKIVGGYTCAANSIPYQVSLNSGSHFCGGSGLINSQWVVSAAHCYKSRIQVRLGE</div><div>708090100110120</div><div>HNIDVLEGNEQFINAAKIITHPNFNGNTLDNDIMLIKLSSPATLNSRVATVSLPRSCAAA</div><div>130140150160170180</div><div>GTECLISGWGNTKSSGSSYP SLLQCLKAPV LSDSSCKSSYPGQITGNMIC VGFLEGGKDS</div><div>190200210220230</div><div>CQGDSGGPVV CNGQLQGIVS WYGCAQKNK PGVYTKVCNY VNWIQQTIAA N</div></div>		

« Hide

References

[1]	"On trypsinogen and trypsin of pig." Charles M., Rovey M., Guidoni A.A., Desnuelle P.
-----	--



	Biochim. Biophys. Acta 69:115-129(1963) [PubMed: 14020231] [Abstract] <u>Cited for:</u> PROTEIN SEQUENCE OF 1-10.
[2]	<b>"Determination of the amino acid sequence of porcine trypsin by sequenator analysis."</b> Hermodson M.A., Ericsson L.H., Neurath H., Walsh K.A. Biochemistry 12:3146-3153(1973) [PubMed: 4738933] [Abstract] <u>Cited for:</u> PROTEIN SEQUENCE OF 9-231.
[3]	<b>"Refined 1.6-A resolution crystal structure of the complex formed between porcine beta-trypsin and MCTI-A, a trypsin inhibitor of the squash family. Detailed comparison with bovine beta-trypsin and its complex."</b> Huang Q., Liu S., Tang Y. J. Mol. Biol. 229:1022-1036(1993) [PubMed: 8445634] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
[4]	<b>"Amino acid sequencing of a trypsin inhibitor by refined 1.6 A X-ray crystal structure of its complex with porcine beta-trypsin."</b> Huang Q., Liu S., Tang Y., Zeng F., Qian R. FEBS Lett. 297:143-146(1992) [PubMed: 1551419] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
[5]	<b>"Refined 1.8-A resolution crystal structure of the porcine epsilon-trypsin."</b> Huang Q., Wang Z., Li Y., Liu S., Tang Y. Biochim. Biophys. Acta 1209:77-82(1994) [PubMed: 7947985] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
[6]	<b>"The three-dimensional structure of recombinant leech-derived tryptase inhibitor in complex with trypsin. Implications for the structure of human mast cell tryptase and its inhibition."</b> Stubbs M.T., Morenweiser R., Stuerzebecher J., Bauer M., Bode W., Huber R., Piechottka G.P., Matschiner G., Sommerhoff C.P., Fritz H., Auerswald E.A. J. Biol. Chem. 272:19931-19937(1997) [PubMed: 9242660] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF COMPLEX WITH LDTI.
[7]	<b>"Structure of the complex of leech-derived tryptase inhibitor (LDTI) with trypsin and modeling of the LDTI-tryptase system."</b> di Marco S., Priestle J.P. Structure 5:1465-1474(1997) [PubMed: 9384562] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (2.03 ANGSTROMS) OF COMPLEX WITH LDTI.
[8]	<b>"L-isoaspartate 115 of porcine beta-trypsin promotes crystallization of its complex with bdellastasin."</b> Rester U., Moser M., Huber R., Bode W. Acta Crystallogr. D 56:581-588(2000) [PubMed: 10771427] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF COMPLEX WITH BDELLASTASIN.
+	Additional computationally mapped references.

## Cross-references

### Sequence databases

PIR | TRPGTR. A90641.

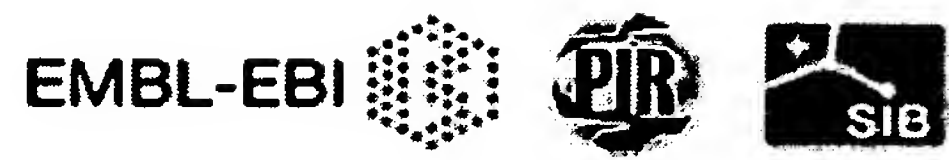
### 3D structure databases

<div> <div>⦿</div> <div>PDBe</div> </div> <div> <div>⦿</div> <div>RCSB PDB</div> </div> <div> <div>⦿</div> <div>PDBj</div> </div>	Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
	1AKS	X-ray	1.80	A	9-133	[>>]
				B	134-231	[>>]

	1AN1	X-ray	2.03	E	9-231	[»]
	1AVW	X-ray	1.75	A	9-231	[»]
	1AVX	X-ray	1.90	A	9-231	[»]
	1C9P	X-ray	2.80	A	9-231	[»]
	1D3O	model	-	A	9-231	[»]
	1DF2	model	-	A	9-231	[»]
	1EJA	X-ray	2.70	A	9-231	[»]
	1EPT	X-ray	1.80	A	9-51	[»]
				B	52-133	[»]
				C	134-231	[»]
	1EWU	model	-	A	9-231	[»]
	1FMG	X-ray	1.90	A	9-231	[»]
	1FN6	X-ray	1.80	A	9-231	[»]
	1FNI	X-ray	1.60	A	9-231	[»]
	1H9H	X-ray	1.50	E	9-231	[»]
	1H9I	X-ray	1.90	E	9-231	[»]
	1LDT	X-ray	1.90	T	9-231	[»]
	1LT2	model	-	A	9-231	[»]
	1MCT	X-ray	1.60	A	9-231	[»]
	1QQU	X-ray	1.63	A	9-231	[»]
	1S5S	X-ray	1.40	A	9-231	[»]
	1S6F	X-ray	1.80	A	9-231	[»]
	1S6H	X-ray	1.45	A	9-231	[»]
	1S81	X-ray	1.70	A	9-231	[»]
	1S82	X-ray	1.85	A	9-231	[»]
	1S83	X-ray	1.25	A	9-231	[»]
	1S84	X-ray	1.85	A	9-231	[»]
	1S85	X-ray	2.20	A	9-231	[»]
	1TFX	X-ray	2.60	A/B	9-231	[»]
	1TX6	X-ray	2.20	A/B/C/D	9-231	[»]
	1UHB	X-ray	2.15	A	9-133	[»]
				B	134-231	[»]
				P	177-185	[»]
	1V6D	X-ray	1.90	A	9-231	[»]
	1YF4	X-ray	1.98	A	9-231	[»]
	1Z7K	X-ray	1.90	A	9-231	[»]
	2A31	X-ray	1.25	A	9-231	[»]
	2A32	X-ray	1.50	A	9-231	[»]
	3MYW	X-ray	2.50	A/B	9-231	[»]
ProteinModelPortal	P00761.					
SMR	P00761. Positions 9-231.					
ModBase	Search...					
Protein-protein interaction databases						
DIP	DIP-6083N.					
Proteomic databases						
PRIDE	P00761.					
Genome annotation databases						
Ensembl	ENSSSCT00000017947; ENSSSCP00000017465; ENSSSCG00000016483.					
Phylogenomic databases						
GeneTree	ENSGT00580000081320.					
HOVERGEN	HBG013304.					

OrthoDB	EOG4SJ5FV.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.21.4. 249.
<b>Family and domain databases</b>	
InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Entry information</b>	
Entry name	TRYP_PIG
Accession	Primary (citable) accession number: <b>P00761</b>
Entry history	Integrated into July 21, 1986 UniProtKB/Swiss-Prot: Last sequence July 21, 1986 update: Last modified: April 5, 2011 This is version 100 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Relevant documents</b>	
PDB cross-references Index of Protein Data Bank (PDB) cross-references	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	





**P07477 (TRY1\_HUMAN) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 133.

**Names and origin**

Protein names	<p><i>Recommended name:</i>  <b>Trypsin-1</b>  EC=3.4.21.4</p> <p><i>Alternative name(s):</i>  Beta-trypsin  Cationic trypsinogen  Serine protease 1  Trypsin I</p> <p><u>Cleaved into the following 2 chains:</u>  1. <b>Alpha-trypsin chain 1</b>  2. <b>Alpha-trypsin chain 2</b></p>
Gene names	<p>Name: <b>PRSS1</b>  Synonyms: TRP1, TRY1, TRYP1</p>
Organism	<b>Homo sapiens (Human)</b> [Complete proteome]
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo

**Protein attributes**

Sequence length	247 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Has activity against the synthetic substrates Boc-Phe-Ser-Arg-Mec, Boc-Leu-Thr-Arg-Mec, Boc-Gln-Ala-Arg-Mec and Boc-Val-Pro-Arg-Mec. The single-chain form is more active than the two-chain form against all of these substrates. <a href="#">Ref.9</a>
Catalytic activity	Preferential cleavage: Arg- -Xaa, Lys- -Xaa.
Cofactor	Binds 1 calcium ion per subunit.
Subcellular location	Secreted › extracellular space.
Post-translational modification	Occurs in a single-chain form and a two-chain form, produced by proteolytic cleavage after Arg-122.
Involvement in disease	Defects in PRSS1 are a cause of hereditary pancreatitis (HPC) [MIM:167800]; also known as chronic pancreatitis (CP). HPC is an autosomal dominant disease characterized by the presence of calculi in pancreatic ducts. It causes severe abdominal pain

	attacks. <span>Ref.8</span> <span>Ref.11</span> <span>Ref.12</span> <span>Ref.16</span> <span>Ref.17</span> <span>Ref.18</span> <span>Ref.19</span> <span>Ref.20</span> <span>Ref.21</span> <span>Ref.22</span> <span>Ref.23</span>
Sequence similarities	Belongs to the peptidase S1 family. Contains 1 peptidase S1 domain.
Caution	Tyr-154 was proposed to be phosphorylated ( <span>Ref.15</span> ) but it has been shown ( <span>Ref.14</span> ) to be sulfated instead. Phosphate and sulfate groups are similar in mass and size, and this can lead to erroneous interpretation of the results.
Mass spectrometry	Molecular mass is 24348±2 Da from positions 24 - 247. Determined by ESI. <span>Ref.15</span>



Ontologies

Keywords

Biological process	Digestion
Cellular component	Secreted
Coding sequence diversity	Polymorphism
Disease	Disease mutation
Domain	Signal
Ligand	Calcium Metal-binding
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Sulfation Zymogen
Technical term	3D-structure Complete proteome Direct protein sequencing

Gene Ontology (GO)

Biological process	digestion Inferred from electronic annotation. Source: UniProtKB-KW  proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	extracellular space Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	metal ion binding Inferred from electronic annotation. Source: UniProtKB-KW  protein binding Inferred from physical interaction. Source: IntAct  serine-type endopeptidase activity Traceable author statement <a href="#">[Ref.11]</a> . Source: UniProtKB

Complete GO annotation...

Binary interactions

With	Entry	#Exp.	IntAct	Notes
GRB2	P62993	1	EBI-1222902,EBI-401755	

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature
-------------	-------------	--------	-------------	----------------	---------

## Molecule processing

	Signal peptide	1 – 15	15	Ref.9 Ref.10		
	Propeptide	16 – 23	8	Activation peptide		PRO_01
	Chain	24 – 247	224	Trypsin-1		PRO_01
	Chain	24 – 122	99	Alpha-trypsin chain 1		PRO_01
	Chain	123 – 247	125	Alpha-trypsin chain 2		PRO_01

## Regions

	Domain	24 – 244	221	Peptidase S1	
--	--------	----------	-----	--------------	--

## Sites

	Active site	63	1	Charge relay system		
	Active site	107	1	Charge relay system		
	Active site	200	1	Charge relay system		
	Metal binding	75	1	Calcium		
	Metal binding	77	1	Calcium; via carbonyl oxygen		
	Metal binding	80	1	Calcium; via carbonyl oxygen		
	Metal binding	85	1	Calcium		
	Site	194	1	Required for specificity By similarity		

## Amino acid modifications

	Modified residue	154	1	Sulfotyrosine Ref.14		
	Disulfide bond	30 ↔ 160				
	Disulfide bond	48 ↔ 64				
	Disulfide bond	139 ↔ 206				
	Disulfide bond	171 ↔ 185				
	Disulfide bond	196 ↔ 220				

## Natural variations

	Natural variant	16	1	A → V in HPC; disrupts signal sequence cleavage site. Ref.18		VAR_01
	Natural variant	22	1	D → G in HPC; increased rate of activation. Ref.8		VAR_01

<input type="checkbox"/>	Natural variant	23	1	K → R in HPC; increased rate of activation. <a href="#">Ref.19</a>	I	VAR_000000
<input type="checkbox"/>	Natural variant	29	1	N → I in HPC. <a href="#">Ref.12</a> <a href="#">Ref.16</a> <a href="#">Ref.17</a> <a href="#">Ref.23</a>	I	VAR_000000
<input type="checkbox"/>	Natural variant	29	1	N → T in HPC. <a href="#">Ref.21</a>	I	VAR_000000
<input type="checkbox"/>	Natural variant	54	1	N → S in HPC; associated with Ile-29; the double mutant shows increased autocatalytic activation which is solely due to the Ile-29 mutation. <a href="#">Ref.23</a>	I	VAR_000000
<input type="checkbox"/>	Natural variant	79	1	E → K in HPC; Lys-79 trypsin activates anionic trypsinogen PRSS2 2-fold while the common pancreatitis-associated mutants His-122 or Ile-29 have no such effect. <a href="#">[dbSNP:rs28934902]</a> <a href="#">Ref.22</a>	I	VAR_000000
<input type="checkbox"/>	Natural variant	104	1	L → P in HPC. <a href="#">Ref.12</a>	I	VAR_000000
<input type="checkbox"/>	Natural variant	116	1	R → C in HPC. <a href="#">Ref.12</a>	I	VAR_000000
<input type="checkbox"/>	Natural variant	122	1	R → C in HPC; suppresses an autocleavage site. <a href="#">Ref.21</a>	I	VAR_000000
<input type="checkbox"/>	Natural variant	122	1	R → H in HPC; suppresses an autocleavage site which is probably part of a fail-safe mechanism by which trypsin, which is activated within the pancreas, may be inactivated; loss of this cleavage site would permit autodigestion resulting in pancreatitis. <a href="#">Ref.11</a> <a href="#">Ref.12</a> <a href="#">Ref.16</a> <a href="#">Ref.18</a> <a href="#">Ref.20</a>	I	VAR_000000



<input type="checkbox"/>	Natural variant	137	1	T → M in a colorectal cancer sample; somatic mutation. <a href="#">Ref.24</a>		VAR_03
<input type="checkbox"/>	Natural variant	139	1	C → F in HPC. <a href="#">Ref.12</a>		VAR_03

Experimental info

<input type="checkbox"/>	Mutagenesis	154	1	Y → F: Lack of sulfation. <a href="#">Ref.14</a>		
<input type="checkbox"/>	Sequence conflict	4	1	L → F in AAI28227. <a href="#">Ref.7</a>		

Secondary structure

1. ....

■ Helix ■ Strand ■ Turn





Details...

Sequences

Sequence	Length	Mass (Da)
<input checked="" type="checkbox"/> P07477 [UniParc]. Last modified April 1, 1988. Version 1. Checksum: DD49A487B8062813	FASTA 247	26,558
<div><div>102030405060 MNPLLLILTFV AAALAAPFDD DDKIVGGYNC EENSVPYQVS LNSGYHFCGG SLINEQWVVS</div><div>708090100110120 AGHCYKSRIQ VRLGEHNIEV LEGNEQFINA AKIIRHPQYD RKTLLNNDIML IKLSSRAVIN</div><div>130140150160170180 ARVSTISLPT APPATGTKCL ISGWGNTASS GADYPDELQC LDAPVLSQAK CEASYPGKIT</div><div>190200210220230240 SNMFCVGFL EGGKDSCQGDS GGPVVCNGQL QGVVSWGDGC AQKNKPGVYT KVINYVKWIK</div><div>NTIAANS</div></div> <div>« Hide</div>		

References

	« Hide 'large scale' references
[1]	"Cloning, characterization and nucleotide sequences of two cDNAs encoding human pancreatic trypsinogens." Emi M., Nakamura Y., Ogawa M., Yamamoto T., Nishide T., Mori T., Matsubara K. Gene 41:305-310(1986) [PubMed: 3011602] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].

[2]	<p><b>"The complete 685-kilobase DNA sequence of the human beta T cell receptor locus."</b>  Rowen L., Koop B.F., Hood L.  Science 272:1755-1762(1996) [PubMed: 8650574] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA].</p>
[3]	<p><b>"Complete sequencing and characterization of 21,243 full-length human cDNAs."</b>  Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.  Sugano S.  Nat. Genet. 36:40-45(2004) [PubMed: 14702039] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  <u>Tissue:</u> Prostate.</p>
[4]	<p><b>"The DNA sequence of human chromosome 7."</b>  Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H., Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R., Wylie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaller M.E., Fewell G.A., Delehaunty K.D., Miner T.L.  Wilson R.K.  Nature 424:157-164(2003) [PubMed: 12853948] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].</p>
[5]	<p><b>"Human chromosome 7: DNA sequence and biology."</b>  Scherer S.W., Cheung J., MacDonald J.R., Osborne L.R., Nakabayashi K., Herbrick J.-A., Carson A.R., Parker-Katiraei L., Skaug J., Khaja R., Zhang J., Hudek A.K., Li M., Haddad M., Duggan G.E., Fernandez B.A., Kanematsu E., Gentles S.  Tsui L.-C.  Science 300:767-772(2003) [PubMed: 12690205] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].</p>
[6]	<p>Mural R.J., Istrail S., Sutton G.G., Florea L., Halpern A.L., Mobarry C.M., Lippert R., Walenz B., Shatkay H., Dew I., Miller J.R., Flanigan M.J., Edwards N.J., Bolanos R., Fasulo D., Halldorsson B.V., Hannenhalli S., Turner R.  Venter J.C.  Submitted (JUL-2005) to the EMBL/GenBank/DDBJ databases  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].</p>
[7]	<p><b>"The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC)."</b>  The MGC Project Team  Genome Res. 14:2121-2127(2004) [PubMed: 15489334] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].</p>
[8]	<p><b>"Chronic pancreatitis associated with an activation peptide mutation that facilitates trypsin activation."</b>  Teich N., Ockenga J., Hoffmeister A., Manns M., Mossner J., Keim V.  Gastroenterology 119:461-465(2000) [PubMed: 10930381] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 15-67, VARIANT HPC GLY-22.</p>
[9]	<p><b>"Identification of one- and two-chain forms of trypsinogen 1 produced by a human gastric adenocarcinoma cell line."</b>  Koshikawa N., Yasumitsu H., Nagashima Y., Umeda M., Miyazaki K.  Biochem. J. 303:187-190(1994) [PubMed: 7945238] [Abstract]  <u>Cited for:</u> PROTEIN SEQUENCE OF 16-43 AND 123-142, FUNCTION, POST-TRANSLATIONAL PROCESSING.  <u>Tissue:</u> Gastric adenocarcinoma.</p>
[10]	<p><b>"Immunoreactive anionic and cationic trypsin in human serum."</b>  Kimland M., Russick C., Marks W.H., Borgstroem A.  Clin. Chim. Acta 184:31-46(1989) [PubMed: 2598466] [Abstract]  <u>Cited for:</u> PROTEIN SEQUENCE OF 16-43.</p>
[11]	<p><b>"Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene."</b>  Whitcomb D.C., Gorry M.C., Preston R.A., Furey W., Sossenheimer M.J., Ulrich C.D., Martin S.P., Gates L.K. Jr., Amann S.T., Toskes P.P., Liddle R., McGrath K., Uomo G., Post J.C., Ehrlich G.D.</p>

	Nat. Genet. 14:141-145(1996) [PubMed: 8841182] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 68-151, VARIANT HPC HIS-122.
[12]	<b>"Mutational screening of patients with nonalcoholic chronic pancreatitis: identification of further trypsinogen variants."</b> Teich N., Bauer N., Mossner J., Keim V. Am. J. Gastroenterol. 97:341-346(2002) [PubMed: 11866271] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 68-151, VARIANTS HPC ILE-29; PRO-104; CYS-116; HIS-122 AND PHE-139.
[13]	Lubec G., Afjehi-Sadat L. Submitted (MAR-2007) to UniProtKB <u>Cited for:</u> PROTEIN SEQUENCE OF 73-92, MASS SPECTROMETRY. <u>Tissue:</u> Brain and Cajal-Retzius cell.
[14]	<b>"Human cationic trypsinogen is sulfated on Tyr154."</b> Sahin-Toth M., Kukor Z., Nemoda Z. FEBS J. 273:5044-5050(2006) [PubMed: 17087724] [Abstract] <u>Cited for:</u> SULFATION AT TYR-154, MUTAGENESIS OF TYR-154.
[15]	<b>"Crystal structure of human trypsin 1: unexpected phosphorylation of Tyr151."</b> Gaboriaud C., Serre L., Guy-Crotte O., Forest E., Fontecilla-Camps J.-C. J. Mol. Biol. 259:995-1010(1996) [PubMed: 8683601] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS), MASS SPECTROMETRY.
[16]	<b>"Mutations in the cationic trypsinogen gene are associated with recurrent acute and chronic pancreatitis."</b> Gorry M.C., Gabbaizedeh D., Furey W., Gates L.K. Jr., Preston R.A., Aston C.E., Zhang Y., Ulrich C., Ehrlich G.D., Whitcomb D.C. Gastroenterology 113:1063-1068(1997) [PubMed: 9322498] [Abstract] <u>Cited for:</u> VARIANTS HPC ILE-29 AND HIS-122.
[17]	<b>"Mutations of the cationic trypsinogen in hereditary pancreatitis."</b> Teich N., Mossner J., Keim V. Hum. Mutat. 12:39-43(1998) [PubMed: 9633818] [Abstract] <u>Cited for:</u> VARIANT HPC ILE-29.
[18]	<b>"A signal peptide cleavage site mutation in the cationic trypsinogen gene is strongly associated with chronic pancreatitis."</b> Witt H., Luck W., Becker M. Gastroenterology 117:7-10(1999) [PubMed: 10381903] [Abstract] <u>Cited for:</u> VARIANTS HPC VAL-16 AND HIS-122.
[19]	<b>"Mutations in the cationic trypsinogen gene and evidence for genetic heterogeneity in hereditary pancreatitis."</b> Ferec C., Ragues O., Salomon R., Roche C., Bernard J.P., Guillot M., Quere I., Faure C., Mercier B., Audrezet M.-P., Guillausseau P.J., Dupont C., Munnich A., Bignon J.D., Le Bodic L. J. Med. Genet. 36:228-232(1999) [PubMed: 10204851] [Abstract] <u>Cited for:</u> VARIANT HPC ARG-23.
[20]	<b>"A CGC&gt;CAT gene conversion-like event resulting in the R122H mutation in the cationic trypsinogen gene and its implication in the genotyping of pancreatitis."</b> Chen J.-M., Ragues O., Ferec C., Deprez P.H., Verellen-Dumoulin C. J. Med. Genet. 37:E36-E36(2000) [PubMed: 11073545] [Abstract] <u>Cited for:</u> VARIANT HPC HIS-122.
[21]	<b>"Novel cationic trypsinogen (PRSS1) N29T and R122C mutations cause autosomal dominant hereditary pancreatitis."</b> Pfutzer R., Myers E., Applebaum-Shapiro S., Finch R., Ellis I., Neoptolemos J., Kant J.A., Whitcomb D.C. Gut 50:271-272(2002) [PubMed: 11788572] [Abstract] <u>Cited for:</u> VARIANTS HPC THR-29 AND CYS-122.



[22]	<p><b>"Interaction between trypsinogen isoforms in genetically determined pancreatitis: mutation E79K in cationic trypsin (PRSS1) causes increased transactivation of anionic trypsinogen (PRSS2)."</b></p> <p>Teich N., Le Marechal C., Kukor Z., Caca K., Witzigmann H., Chen J.-M., Toth M., Moessner J., Keim V., Ferec C., Sahin-Toth M.</p> <p>Hum. Mutat. 23:22-31(2004) [PubMed: 14695529] [Abstract]</p> <p><u>Cited for:</u> VARIANT HPC LYS-79, CHARACTERIZATION OF VARIANT HPC LYS-79.</p>
[23]	<p><b>"Gene conversion between functional trypsinogen genes PRSS1 and PRSS2 associated with chronic pancreatitis in a six-year-old girl."</b></p> <p>Teich N., Nemoda Z., Koehler H., Heinritz W., Moessner J., Keim V., Sahin-Toth M.</p> <p>Hum. Mutat. 25:343-347(2005) [PubMed: 15776435] [Abstract]</p> <p><u>Cited for:</u> VARIANTS HPC ILE-29 AND SER-54, CHARACTERIZATION OF VARIANTS HPC ILE-29 AND SER-54.</p>
[24]	<p><b>"The consensus coding sequences of human breast and colorectal cancers."</b></p> <p>Sjoeblom T., Jones S., Wood L.D., Parsons D.W., Lin J., Barber T.D., Mandelker D., Leary R.J., Ptak J., Silliman N., Szabo S., Buckhaults P., Farrell C., Meeh P., Markowitz S.D., Willis J., Dawson D., Willson J.K.V., Velculescu V.E.</p> <p>Science 314:268-274(2006) [PubMed: 16959974] [Abstract]</p> <p><u>Cited for:</u> VARIANT [LARGE SCALE ANALYSIS] MET-137.</p>
+	Additional computationally mapped references.

Web resources

GeneReviews

Cross-references

Sequence databases

EMBL	M22612 mRNA. Translation: AAA61231.1.
GenBank	L36092 Genomic DNA. Translation: AAC80207.1.
DDBJ	AK312199 mRNA. Translation: BAG35132.1.
	AC231380 Genomic DNA. No translation available.
	CH236959 Genomic DNA. Translation: EAL23773.1.
	CH471198 Genomic DNA. Translation: EAW51925.1.
	BC128226 mRNA. Translation: AAI28227.1.
	AF314534 Genomic DNA. Translation: AAG30943.1.
	U70137 Genomic DNA. Translation: AAC50728.1.
	AF315309 Genomic DNA. Translation: AAG30947.1.
	AF315310 Genomic DNA. Translation: AAG30948.1.
	AF315311 Genomic DNA. Translation: AAG30949.1.

IPI	IPI00011694.
PIR	A25852. S50020. S50021.
RefSeq	NP_002760.1. NM_002769.4.
UniGene	Hs.449281.

3D structure databases

PDBe	Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
RCSB PDB	1FXY	X-ray	2.15	A	127-247	[»]
PDBj	1TRN	X-ray	2.20	A/B	24-247	[»]

	2RA3 X-ray 1.46 A/B 24-247 [»]
ProteinModelPortal	P07477.
SMR	P07477. Positions 24-247.
ModBase	Search...
<b>Protein-protein interaction databases</b>	
IntAct	P07477. 3 interactions.
STRING	P07477.
<b>Protein family/group databases</b>	
MEROPS	S01.127.
<b>PTM databases</b>	
PhosphoSite	P07477.
<b>Proteomic databases</b>	
PeptideAtlas	P07477.
PRIDE	P07477.
<b>Genome annotation databases</b>	
Ensembl	ENST00000311737; ENSP00000308720; ENSG00000204983.
GeneID	5644.
KEGG	hsa:5644.
UCSC	uc003wak.2. human.
<b>Organism-specific databases</b>	
CTD	5644.
GeneCards	GC07P142458.
H-InvDB	HIX0058295.
HGNC	HGNC:9475. PRSS1.
HPA	CAB025487. CAB025538.
MIM	167800. phenotype. 276000. gene+phenotype.
neXtProt	NX_P07477.
Orphanet	676. Hereditary chronic pancreatitis.
PharmGKB	PA33828.
GenAtlas	Search...
<b>Phylogenomic databases</b>	
HOVERGEN	HBG013304.
InParanoid	P07477.
OMA	WIKNTIA.

**Enzyme and pathway databases**

BRENDA	3.4.21.4. 247.
--------	----------------

**Gene expression databases**

ArrayExpress	P07477.
Bgee	P07477.
Genevestigator	P07477.
GermOnline	ENSG00000173636. Homo sapiens.


**Family and domain databases**

InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...

**Other Resources**

NextBio	21926.
PMAP-CutDB	P07477.
SOURCE	Search...

**Entry information**

Entry name	TRY1_HUMAN
Accession	Primary (citable) accession number: <b>P07477</b> Secondary accession number(s): A1A509  Q9HAN7
Entry history	Integrated into April 1, 1988 UniProtKB/Swiss-Prot: Last sequence April 1, 1988 update: Last modified: April 5, 2011 This is version 133 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program



**Disclaimer**

Any medical or genetic information present in this entry is provided for research, educational and informational purposes only. It is not in any way intended to be used as a substitute for professional medical advice, diagnosis, treatment or care.

**Relevant documents**

Human chromosome 7

Human chromosome 7: entries, gene names and cross-references to MIM

Human entries with polymorphisms or disease mutations

List of human entries with polymorphisms or disease mutations

Human polymorphisms and disease mutations

Index of human polymorphisms and disease mutations

MIM cross-references

Online Mendelian Inheritance in Man (MIM) cross-references in UniProtKB/Swiss-Prot

PDB cross-references

Index of Protein Data Bank (PDB) cross-references

Peptidase families

Classification of peptidase families and list of entries

SIMILARITY comments

Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact

EMBL-EBI



**P07478 (TRY2\_HUMAN) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 116.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Trypsin-2</b> EC=3.4.21.4 <i>Alternative name(s):</i> Anionic trypsinogen Serine protease 2 Trypsin II
Gene names	Name: <b>PRSS2</b> Synonyms: TRY2, TRYP2
Organism	<b>Homo sapiens (Human)</b> [Complete proteome]
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo

**Protein attributes**

Sequence length	247 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Catalytic activity	Preferential cleavage: Arg- -Xaa, Lys- -Xaa.
Cofactor	Binds 1 calcium ion per subunit <a href="#">[By similarity]</a> .
Subcellular location	Secreted › extracellular space.
Post-translational modification	Sulfated on tyrosine.
Sequence similarities	Belongs to the peptidase S1 family. Contains 1 peptidase S1 domain.





**Ontologies****Keywords**

Biological process	Digestion
Cellular component	Secreted
Coding sequence diversity	Polymorphism
Domain	Signal

Ligand	Calcium Metal-binding
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Phosphoprotein Sulfation Zymogen
Technical term	Complete proteome Direct protein sequencing
Gene Ontology (GO)	
Biological process	collagen catabolic process Inferred from direct assay. Source: UniProtKB  digestion Inferred from electronic annotation. Source: UniProtKB-KW  positive regulation of cell adhesion Traceable author statement. Source: UniProtKB  positive regulation of cell growth Traceable author statement. Source: UniProtKB  proteolysis Inferred from direct assay. Source: UniProtKB
Cellular component	extracellular matrix Traceable author statement. Source: UniProtKB  extracellular space Inferred from mutant phenotype. Source: UniProtKB
Molecular function	calcium ion binding Inferred from direct assay. Source: UniProtKB  protein binding Inferred from physical interaction. Source: UniProtKB  serine-type endopeptidase activity Inferred from direct assay. Source: UniProtKB

Complete GO annotation...

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature i
Molecule processing					
	Signal peptide	1 – 15	15		
	Propeptide	16 – 23	8	Activation peptide	PRO_001
	Chain	24 – 247	224	Trypsin-2	PRO_001



Regions

	Domain	24 – 244	221	Peptidase S1		
---	--------	----------	-----	--------------	--	--

Sites

<input type="checkbox"/>	Active site	63	1	Charge relay system By similarity		
<input type="checkbox"/>	Active site	107	1	Charge relay system By similarity		
<input type="checkbox"/>	Active site	200	1	Charge relay system By similarity		
<input type="checkbox"/>	Metal binding	75	1	Calcium By similarity		
<input type="checkbox"/>	Metal binding	77	1	Calcium; via carbonyl oxygen By similarity		
<input type="checkbox"/>	Metal binding	80	1	Calcium; via carbonyl oxygen By similarity		
<input type="checkbox"/>	Metal binding	85	1	Calcium By similarity		
<input type="checkbox"/>	Site	194	1	Required for specificity By similarity		


Amino acid modifications

<input type="checkbox"/>	Modified residue	8	1	Phosphothreonine Ref.4		
<input type="checkbox"/>	Modified residue	154	1	Sulftirosine Probable		
<input type="checkbox"/>	Disulfide bond	30 ↔ 160		By similarity		
<input type="checkbox"/>	Disulfide bond	48 ↔ 64		By similarity		
<input type="checkbox"/>	Disulfide bond	171 ↔ 185		By similarity		
<input type="checkbox"/>	Disulfide bond	196 ↔ 220		By similarity		

Natural variations

<input type="checkbox"/>	Natural variant	117	1	A → V. [dbSNP:rs11547028]		VAR_05
--------------------------	-----------------	-----	---	------------------------------	--	--------

Sequences

Sequence		Length	Mass (Da)
	P07478 [UniParc]. Last modified April 1, 1988. Version 1. Checksum: 82B0F41EB8E3D5DB	FASTA 247	26,488

<u>10</u>	<u>20</u>	<u>30</u>	<u>40</u>	<u>50</u>	<u>60</u>
MNLLLILTFV	AAAVAAPFDD	DDKIVGGYIC	EENSVPYQVS	LNSGYHFCGG	SLISEQWVVS
<u>70</u>	<u>80</u>	<u>90</u>	<u>100</u>	<u>110</u>	<u>120</u>
AGHCYKSRIQ	VRIGEHNIEV	LEGNEQFINA	AKIIRHPKYN	SRTLDNDILL	IKLSSPAVIN
<u>130</u>	<u>140</u>	<u>150</u>	<u>160</u>	<u>170</u>	<u>180</u>
SRVSAISLPT	APPAAGTESL	ISGWGNTLSS	GADYPDELQC	LDAPVLSQAE	CEASYPGKIT
<u>190</u>	<u>200</u>	<u>210</u>	<u>220</u>	<u>230</u>	<u>240</u>
NNMFCVGFLE	GGKDSCQGDS	GGPVVSNGL	QGIVSWG YGC	AQKNRPGVYT	KVYNYVDWIK
DTIAANS					

« Hide

References

	« Hide 'large scale' references
[1]	<b>"Cloning, characterization and nucleotide sequences of two cDNAs encoding human pancreatic trypsinogens."</b> Emi M., Nakamura Y., Ogawa M., Yamamoto T., Nishide T., Mori T., Matsubara K. Gene 41:305-310(1986) [PubMed: 3011602] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].
[2]	<b>"Immunoreactive anionic and cationic trypsin in human serum."</b> Kimland M., Russick C., Marks W.H., Borgstroem A. Clin. Chim. Acta 184:31-46(1989) [PubMed: 2598466] [Abstract] <u>Cited for:</u> PROTEIN SEQUENCE OF 16-49.
[3]	<b>"Human cationic trypsinogen is sulfated on Tyr154."</b> Sahin-Toth M., Kukor Z., Nemoda Z. FEBS J. 273:5044-5050(2006) [PubMed: 17087724] [Abstract] <u>Cited for:</u> SULFATION.
[4]	<b>"Automated phosphoproteome analysis for cultured cancer cells by two-dimensional nanoLC-MS using a calcined titania/C18 biphasic column."</b> Imami K., Sugiyama N., Kyono Y., Tomita M., Ishihama Y. Anal. Sci. 24:161-166(2008) [PubMed: 18187866] [Abstract] <u>Cited for:</u> PHOSPHORYLATION [LARGE SCALE ANALYSIS] AT THR-8, MASS SPECTROMETRY. <u>Tissue:</u> Cervix carcinoma.
+	Additional computationally mapped references.

Cross-references

Sequence databases

<input checked="" type="radio"/> EMBL	M27602 mRNA. Translation: AAA61232.1.
<input checked="" type="radio"/> GenBank	
<input checked="" type="radio"/> DDBJ	
IPI	IPI00815665.
PIR	B25852.
RefSeq	NP_002761.1. NM_002770.2.
UniGene	Hs.449281.

**3D structure databases**

ProteinModelPortal	P07478.
SMR	P07478. Positions 24-247.
ModBase	Search...

**Protein-protein interaction databases**

STRING	P07478.
--------	---------

**Protein family/group databases**

MEROPS	S01.258.
--------	----------

**PTM databases**

PhosphoSite	P07478.
-------------	---------

**Proteomic databases**

PRIDE	P07478.
-------	---------

**Genome annotation databases**

Ensembl	ENST00000438955; ENSP00000414534; ENSG00000204983.
GeneID	5645.
KEGG	hsa:5645.
UCSC	uc003wap.1. human.

**Organism-specific databases**

CTD	5645.
GeneCards	GC07P142298.
H-InvDB	HIX0175861.
HGNC	HGNC:9483. PRSS2.
HPA	CAB025487. CAB025538.
MIM	601564. gene.
neXtProt	NX_P07478.
Orphanet	676. Hereditary chronic pancreatitis.
PharmGKB	PA33833.
GenAtlas	Search...

**Phylogenomic databases**

eggNOG	prNOG14047.
HOVERGEN	HBG013304.
InParanoid	P07478.



**Enzyme and pathway databases**

BRENDA	3.4.21.4. 247.
--------	----------------

**Gene expression databases**

ArrayExpress	P07478.
Bgee	P07478.
CleanEx	HS_PRSS2.
Genevestigator	P07478.
GermOnline	ENSG00000204982. Homo sapiens.

**Family and domain databases**

InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...

**Other Resources**

NextBio	21930.
SOURCE	Search...

**Entry information**

Entry name	TRY2_HUMAN
Accession	Primary (citable) accession number: <b>P07478</b>
Entry history	Integrated into April 1, 1988 UniProtKB/Swiss-Prot: Last sequence update: April 1, 1988 Last modified: April 5, 2011 This is version 116 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
Disclaimer	Any medical or genetic information present in this entry is provided for research, educational and informational purposes

only. It is not in any way intended to be used as a substitute for professional medical advice, diagnosis, treatment or care.

### Relevant documents

Human chromosome 7

Human chromosome 7: entries, gene names and cross-references to MIM

Human entries with polymorphisms or disease mutations

List of human entries with polymorphisms or disease mutations

Human polymorphisms and disease mutations

Index of human polymorphisms and disease mutations

MIM cross-references

Online Mendelian Inheritance in Man (MIM) cross-references in UniProtKB/Swiss-Prot

Peptidase families

Classification of peptidase families and list of entries

SIMILARITY comments

Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact

EMBL-EBI



**P35030 (TRY3\_HUMAN) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 114.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Trypsin-3</b> EC=3.4.21.4 <i>Alternative name(s):</i> Brain trypsinogen Mesotrypsinogen Serine protease 3 Serine protease 4 Trypsin III Trypsin IV
Gene names	Name: <b>PRSS3</b> Synonyms: PRSS4, TRY3, TRY4
Organism	<b>Homo sapiens (Human)</b> [Complete proteome]
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo

**Protein attributes**

Sequence length	304 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Digestive protease specialized for the degradation of trypsin inhibitors. <a href="#">Ref.5</a>
Catalytic activity	Preferential cleavage: Arg- -Xaa, Lys- -Xaa.
Cofactor	Binds 1 calcium ion per subunit.
Subcellular location	Secreted.
Tissue specificity	Pancreas and brain.
Sequence similarities	Belongs to the peptidase S1 family. Contains 1 peptidase S1 domain.

**Ontologies****Keywords**



Biological process	Digestion
Cellular component	Secreted
Coding sequence diversity	Alternative splicing Polymorphism
Domain	Signal
Ligand	Calcium Metal-binding
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Sulfation Zymogen
Technical term	3D-structure Complete proteome
<b>Gene Ontology (GO)</b>	
Biological process	digestion Traceable author statement <a href="#">Ref.5</a> . Source: UniProtKB endothelial cell migration Inferred from mutant phenotype. Source: UniProtKB zymogen activation Inferred from direct assay. Source: UniProtKB
Cellular component	extracellular space Inferred from direct assay. Source: UniProtKB
Molecular function	calcium ion binding Inferred from direct assay. Source: UniProtKB protein binding Inferred from physical interaction <a href="#">Ref.5</a> . Source: UniProtKB serine-type endopeptidase activity Inferred from direct assay <a href="#">Ref.5</a> . Source: UniProtKB serine-type peptidase activity Inferred from direct assay. Source: UniProtKB
Complete GO annotation...	
<b>Alternative products</b>	
This entry describes 3 isoforms produced by <b>alternative splicing</b> . <a href="#">[Align]</a> <a href="#">[Select]</a>	
<b>Isoform A (identifier: P35030-1)</b>	
<i>This isoform has been chosen as the 'canonical' sequence. All positional information in this entry refers to it. This is also the sequence that appears in the downloadable versions of the entry.</i>	
<b>Isoform B (identifier: P35030-2)</b>	

The sequence of this isoform differs from the canonical sequence as follows:  
1-45: MCGPDDRCPARWPGPGRAVKCGKGLAAARPGRVERGGAQRGGAGL → M

Isoform C (identifier: P35030-3)

The sequence of this isoform differs from the canonical sequence as follows:  
1-70: MCGPDDRCPA...DADGCEALGT → MNPFLILAFVGAA

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Featu
Molecule processing					
<input type="checkbox"/> Signal peptide	1 – ?		Potential		
<input type="checkbox"/> Propeptide	? – 80		Activation peptide		PRO_
<input checked="" type="checkbox"/> Chain	81 – 304	224	Trypsin-3		PRO_
Regions					
<input checked="" type="checkbox"/> Domain	81 – 301	221	Peptidase S1		
Sites					
<input type="checkbox"/> Active site	120	1	Charge relay system		
<input type="checkbox"/> Active site	164	1	Charge relay system		
<input type="checkbox"/> Active site	257	1	Charge relay system		
<input type="checkbox"/> Metal binding	132	1	Calcium		
<input type="checkbox"/> Metal binding	134	1	Calcium; via carbonyl oxygen		
<input type="checkbox"/> Metal binding	137	1	Calcium; via carbonyl oxygen		
<input type="checkbox"/> Metal binding	142	1	Calcium		
<input type="checkbox"/> Site	251	1	Required for specificity By similarity		
Amino acid modifications					
<input type="checkbox"/> Modified residue	211	1	Sulftirosine By similarity		
<input type="checkbox"/> Disulfide bond	87 ↔ 217				
<input type="checkbox"/> Disulfide bond	105 ↔ 121				
<input type="checkbox"/> Disulfide bond	196 ↔ 263				
<input type="checkbox"/> Disulfide bond	228 ↔ 242				





FCVGFLEGGK DSCQRDSGGP VVCNGQLQGV VSWGHCAGWK NRPGVYTKVY NYVDWIKDTI


AANS

« Hide

<input checked="" type="checkbox"/>	<b>Isoform B.</b>	FASTA	260	28,161
	Checksum: CD8AA6E8072BCE56 Show »			
<input checked="" type="checkbox"/>	<b>Isoform C.</b>	FASTA	247	26,727
	Checksum: 01563656780A6607 Show »			

## References

« Hide 'large scale' references

- [1] **"Cloning of the cDNA encoding human brain trypsinogen and characterization of its product."**  
Wiegand U., Corbach S., Minn A., Kang J., Mueller-Hill B.  
Gene 136:167-175(1993) [PubMed: 8294000] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA] (ISOFORMS A AND B), VARIANT ALA-188.  
Tissue: Brain.
- [2] **"Nucleotide sequence of the human pancreatic trypsinogen III cDNA."**  
Tani T., Kawashima I., Mita K., Takiguchi Y.  
Nucleic Acids Res. 18:1631-1631(1990) [PubMed: 2326201] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM C), VARIANTS ALA-188 AND CYS-232.  
Tissue: Pancreas.
- [3] Fukuoka S.  
Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases  
Cited for: NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM C), VARIANT ALA-188.
- [4] **"DNA sequence and analysis of human chromosome 9."**  
Humphray S.J., Oliver K., Hunt A.R., Plumb R.W., Loveland J.E., Howe K.L., Andrews T.D., Searle S., Hunt S.E., Scott C.E., Jones M.C., Ainscough R., Almeida J.P., Ambrose K.D., Ashwell R.I.S., Babbage A.K., Babbage S., Bagguley C.L.  Dunham I.  
Nature 429:369-374(2004) [PubMed: 15164053] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
- [5] **"Human mesotrypsin is a unique digestive protease specialized for the degradation of trypsin inhibitors."**  
Szmola R., Kukor Z., Sahin-Toth M.  
J. Biol. Chem. 278:48580-48589(2003) [PubMed: 14507909] [Abstract]  
Cited for: FUNCTION.
- [6] **"Crystal structure reveals basis for the inhibitor resistance of human brain trypsin."**  
Katona G., Berglund G.I., Hajdu J., Graf L., Szilagyi L.  
J. Mol. Biol. 315:1209-1218(2002) [PubMed: 11827488] [Abstract]  
Cited for: X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) (ISOFORM A).
- + Additional computationally mapped references.

**Cross-references****Sequence databases**

<ul style="list-style-type: none"> <li>⊙ EMBL</li> <li>⊙ GenBank</li> <li>⊙ DDBJ</li> </ul>	X72781 mRNA. Translation: CAB58178.1. X71345 mRNA. Translation: CAA50484.1. X15505 mRNA. Translation: CAA33527.1. D45417 mRNA. Translation: BAA08257.1. AL356489, AL139113, AL358573 Genomic DNA. Translation: CAH69873.1. AL358573, AL139113 Genomic DNA. Translation: CAI39514.1. AL358573, AL139113, AL356489 Genomic DNA. Translation: CAI39515.1. AL139113, AL358573 Genomic DNA. Translation: CAI39655.1. AL139113, AL356489, AL358573 Genomic DNA. Translation: CAI39658.1.
---	--

IPI	IPI00015614. IPI00220839. IPI00843764.
-----	--

PIR	S12764. S33496.
-----	--------------------

RefSeq	NP_002762.2. NM_002771.3.
--------	---------------------------

UniGene	Hs.654513.
---------	------------

**3D structure databases**

<ul style="list-style-type: none"> <li>⊙ PDBe</li> <li>⊙ RCSB PDB</li> <li>⊙ PDBj</li> </ul>	Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
	1H4W	X-ray	1.70	A	81-304	[»]
	2R9P	X-ray	1.40	A/B/C/D	81-304	[»]
	3L33	X-ray	2.48	A/B/C/D	81-304	[»]

ProteinModelPortal	P35030.
--------------------	---------

SMR	P35030. Positions 81-304.
-----	---------------------------

ModBase	Search...
---------	-----------

**Protein-protein interaction databases**

STRING	P35030.
--------	---------

**Protein family/group databases**

MEROPS	S01.174.
--------	----------

**Proteomic databases**

PRIDE	P35030.
-------	---------

**Genome annotation databases**

Ensembl	ENST00000361005; ENSP00000354280; ENSG00000010438.
---------	--

GeneID	5646.
--------	-------

KEGG	hsa:5646.
------	-----------

UCSC	uc003ztj.2. human.
------	--------------------

**Organism-specific databases**

CTD	5646.
GeneCards	GC09P033708.
HGNC	HGNC:9486. PRSS3.
MIM	613578. gene.
neXtProt	NX_P35030.
GenAtlas	Search...

**Phylogenomic databases**

eggNOG	maNOG08762.
HOGENOM	HBG755338.
HOVERGEN	HBG013304.
InParanoid	P35030.
OrthoDB	EOG4SJ5FV.
PhylomeDB	P35030.

**Enzyme and pathway databases**

BRENDA	3.4.21.4. 247.
--------	----------------

**Gene expression databases**

ArrayExpress	P35030.
Bgee	P35030.
CleanEx	HS_PRSS3.
Genevestigator	P35030.
GermOnline	ENSG00000010438. Homo sapiens.

**Family and domain databases**


InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. False negative. [Graphical view]
ProtoNet	Search...



**Other Resources**

NextBio	21934.
SOURCE	Search...

**Entry information**

Entry name	TRY3_HUMAN
Accession	Primary (citable) accession number: <b>P35030</b> Secondary accession number(s): A9Z1Y4  Q9UQV3
Entry history	Integrated into February 1, 1994 UniProtKB/Swiss-Prot: Last sequence update: October 14, 2008 Last modified: April 5, 2011 This is version 114 of the entry and version 2 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
Disclaimer	Any medical or genetic information present in this entry is provided for research, educational and informational purposes only. It is not in any way intended to be used as a substitute for professional medical advice, diagnosis, treatment or care.

**Relevant documents**

Human chromosome 9 Human chromosome 9: entries, gene names and cross-references to MIM
Human entries with polymorphisms or disease mutations List of human entries with polymorphisms or disease mutations
Human polymorphisms and disease mutations Index of human polymorphisms and disease mutations
MIM cross-references Online Mendelian Inheritance in Man (MIM) cross-references in UniProtKB/Swiss-Prot
PDB cross-references Index of Protein Data Bank (PDB) cross-references
Peptidase families Classification of peptidase families and list of entries
SIMILARITY comments Index of protein domains and families

© 2002–2011 UniProt Consortium | License &amp; Disclaimer | Contact



## Protein

Translations of Life

Display Settings: GenPept

## trypsinogen [Homo sapiens]

GenBank: AAA61231.1

[FASTA](#) [Graphics](#)Go to:

LOCUS AAA61231 247 aa linear PRI 14-JAN-1995  
DEFINITION trypsinogen [Homo sapiens].  
ACCESSION AAA61231  
VERSION AAA61231.1 GI:521216  
DBSOURCE locus HUMTRPSGNA accession M22612.1  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 247)  
AUTHORS Emi,M., Nakamura,Y., Ogawa,M., Yamamoto,T., Nishide,T., Mori,T. and  
Matsubara,K.  
TITLE Cloning, characterization and nucleotide sequences of two cDNAs  
encoding human pancreatic trypsinogens  
JOURNAL Gene 41 (2-3), 305-310 (1986)  
PUBMED 3011602  
COMMENT Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..247  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="7q32-qter"  
/tissue\_type="pancreas"  
Protein 1..247  
/product="trypsinogen"  
sig\_peptide 1..15  
/gene="TRY1"  
/note="G00-119-620"  
mat\_peptide 24..247  
/gene="TRY1"  
/product="trypsinogen"  
/note="G00-119-620"  
Region 24..242  
/region\_name="Tryp\_SpC"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cd00190"  
/db\_xref="CDD:29152"  
Site 24  
/site\_type="cleavage"  
/db\_xref="CDD:29152"  
Site order(63,107,200)  
/site\_type="active"  
/db\_xref="CDD:29152"  
Site order(194,215,217)  
/site\_type="other"  
/note="substrate binding sites"  
/db\_xref="CDD:29152"  
CDS 1..247  
/gene="TRY1"  
/coded\_by="M22612.1:7..750"  
/db\_xref="GDB:G00-119-620"  
ORIGIN  
1 mnplliltfv aaalaapfdd ddkivggync eensvpyqvs lnsghfcgg slineqwwvs  
61 aghcyksriq vrlgehnief legneqfina akiirhpgyd rktlnndiml iklssravin  
121 arvstislpt appatgtkcl isgwgntass gadypdelqc ldapvlsqak ceasypgkit  
181 snmfvcvgle ggdscqgds ggpvcngql qgvvswgdgc aqknkpgvyt kvynyvkwik  
241 ntiaans  
//

## Protein

Translations of Life

Display Settings: GenPept

## cationic trypsinogen [Homo sapiens]

GenBank: AAG30943.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AF314534\_1 53 aa linear PRI 08-NOV-2000  
DEFINITION cationic trypsinogen [Homo sapiens].  
ACCESSION AAG30943  
VERSION AAG30943.1 GI:11120616  
DBSOURCE accession AF314534.1  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 53)  
AUTHORS Teich,N., Ockenga,J., Hoffmeister,A., Manns,M., Mossner,J. and  
Keim,V.  
TITLE Chronic pancreatitis associated with an activation peptide mutation  
that facilitates trypsin activation  
JOURNAL Gastroenterology 119 (2), 461-465 (2000)  
PUBMED 10930381  
REFERENCE 2 (residues 1 to 53)  
AUTHORS Teich,N., Ockenga,J., Hoffmeister,A., Manns,M., Mossner,J. and  
Keim,V.  
TITLE Direct Submission  
JOURNAL Submitted (18-OCT-2000) Medizinische Klinik und Poliklinik II,  
Universitätsklinikum Leipzig, Philipp-Rosenthal-Str. 27, Leipzig  
04103, Germany  
COMMENT Method: conceptual translation supplied by author.  
FEATURES  
Location/Qualifiers  
source 1..53  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="7"  
/map="7q35"  
Protein <1..>53  
/product="cationic trypsinogen"  
/EC\_number="3.4.21.4"  
/name="peptidase"  
Region 10..>53  
/region\_name="Tryp\_SpC"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; c100149"  
/db\_xref="CDD:153545"  
Site 10  
/site\_type="cleavage"  
/db\_xref="CDD:29152"  
CDS 1..53  
/gene="TRYP1"  
/allele="D22G"  
/coded\_by="AF314534.1:<102..>261"  
/codon\_start=3  
ORIGIN  
1 aapfdddgki vggynceens vpyqvslnsg yhfccgslin eqwvvsaghc yks  
//



## Protein

Translations of Life

Display Settings: GenPept

## cationic trypsinogen [Homo sapiens]

GenBank: AAC50728.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAC50728 84 aa linear PRI 08-OCT-1996  
DEFINITION cationic trypsinogen [Homo sapiens].  
ACCESSION AAC50728  
VERSION AAC50728.1 GI:1616766  
DBSOURCE locus HSU70137 accession U70137.1  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 84)  
AUTHORS Whitcomb,D.C., Gorry,M.C., Preston,R.A., Furey,W.,  
Sossenheimer,M.J., Ulrich,C.D., Martin,S.P., Gates,L., Amann,S.T.,  
Toskes,P.P., Liddle,R., McGrath,K., Uomo,G., Post,J.C. and  
Ehrlich,G.D.  
TITLE Hereditary pancreatitis is caused by a mutation in the cationic  
trypsinogen gene  
JOURNAL Nat. Genet. 14 (2), 141-145 (1996)  
PUBMED 8841182  
REFERENCE 2 (residues 1 to 84)  
AUTHORS Whitcomb,D.C., Gorry,M.C., Preston,R.A., Furey,W.,  
Sossenheimer,M.J., Ulrich,C.D., Martin,S.P., Gates,L.K.,  
Amann,S.T., Toskes,P.P., Liddle,R., McGrath,K., Uomo,G., Post,J.C.  
and Ehrlich,G.D.  
TITLE Direct Submission  
JOURNAL Submitted (09-SEP-1996) Pathology, University of Pittsburgh, S-792  
Scaife Hall, Pittsburgh, PA 15261, USA  
FEATURES  
source Location/Qualifiers  
1..84  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="7"  
/map="7q35"  
Protein 1..84  
/product="cationic trypsinogen"  
Region <1..>84  
/region\_name="Tryp\_SpC"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cl00149"  
/db\_xref="CDD:153545"  
CDS 1..84  
/gene="TRYP1"  
/coded\_by="U70137.1:<186..>439"  
/codon\_start=2  
ORIGIN  
1 riqvrlgehn ievlegneqf inaakiirhp qydrktlnnd imliklssra vinahvstis  
61 lptappatgt kclisgwnt assg  
//

## Protein

Translations of Life

Display Settings: GenPept

## trypsinogen [Homo sapiens]

GenBank: AAA61232.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAA61232 247 aa linear PRI 14-JAN-1995  
DEFINITION trypsinogen [Homo sapiens].  
ACCESSION AAA61232  
VERSION AAA61232.1 GI:521218  
DBSOURCE locus HUMTRPSGNB accession M27602.1  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 247)  
AUTHORS Emi,M., Nakamura,Y., Ogawa,M., Yamamoto,T., Nishide,T., Mori,T. and  
Matsubara,K.  
TITLE Cloning, characterization and nucleotide sequences of two cDNAs  
encoding human pancreatic trypsinogens  
JOURNAL Gene 41 (2-3), 305-310 (1986)  
PUBMED 3011602  
COMMENT Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..247  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tissue\_type="pancreas"  
Protein 1..247  
/product="trypsinogen"  
sig\_peptide 1..15  
/gene="TRY2"  
mat\_peptide 24..247  
/gene="TRY2"  
/product="trypsinogen"  
Region 24..242  
/region\_name="Tryp\_SpC"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cd00190"  
/db\_xref="CDD:29152"  
Site 24  
/site\_type="cleavage"  
/db\_xref="CDD:29152"  
Site order(63,107,200)  
/site\_type="active"  
/db\_xref="CDD:29152"  
Site order(194,215,217)  
/site\_type="other"  
/note="substrate binding sites"  
/db\_xref="CDD:29152"  
CDS 1..247  
/gene="TRY2"  
/coded\_by="M27602.1:7..750"  
ORIGIN  
1 mnllliltfv aaavaapfdd ddktivggyic eensvpyqvs lnsghfcgg sliseqwwvs  
61 aghcyksriq vrlgehnief legneqfina akiirhpkyn srtldndill iklsspavin  
121 srvsaislpt appaagtesl isgwntlss gadypdelqc ldapvlsqae ceasypgkit  
181 nnmfcvgfle ggdscqgds ggpvvngel qgivswgygc aqknrpgvyt kvynyvdwik  
241 dtiaans  
//

**P00775 (TRYP\_STRGR) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified January 11, 2011. Version 80.

Names and origin	
Protein names	<i>Recommended name:</i> <b>Trypsin</b> EC=3.4.21.4 <i>Alternative name(s):</i> SGT
Gene names	Name: <b>sprT</b>
Organism	<b>Streptomyces griseus</b>
Taxonomic identifier	1911 [NCBI]
Taxonomic lineage	Bacteria › Actinobacteria › Actinobacteridae › Actinomycetales › Streptomycineae › Streptomycetaceae › Streptomyces
Protein attributes	
Sequence length	259 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.
General annotation (Comments)	
Catalytic activity	Preferential cleavage: Arg- -Xaa, Lys- -Xaa.
Sequence similarities	Belongs to the peptidase S1 family. Contains 1 peptidase S1 domain.
Ontologies	
Keywords	
Domain	Signal
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Zymogen
Technical term	3D-structure Direct protein sequencing
Gene Ontology (GO)	
Biological process	proteolysis Inferred from electronic annotation. Source: InterPro



## Molecular function

serine-type endopeptidase activity  
Inferred from electronic annotation. Source: InterPro

Complete GO annotation...

## Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
<b>Molecule processing</b>					
<input checked="" type="checkbox"/> Signal peptide	1 – 32	32			
<input type="checkbox"/> Propeptide	33 – 36	4	Activation peptide		PRO_000002830
<input checked="" type="checkbox"/> Chain	37 – 259	223	Trypsin		PRO_000002830
<b>Regions</b>					
<input checked="" type="checkbox"/> Domain	37 – 257	221	Peptidase S1		
<b>Sites</b>					
<input type="checkbox"/> Active site	73	1	Charge relay system		
<input type="checkbox"/> Active site	118	1	Charge relay system		
<input type="checkbox"/> Active site	208	1	Charge relay system		
<input type="checkbox"/> Site	202	1	Required for specificity		
<b>Amino acid modifications</b>					
<input type="checkbox"/> Disulfide bond	58 ↔ 74				
<input type="checkbox"/> Disulfide bond	177 ↔ 192				
<input type="checkbox"/> Disulfide bond	204 ↔ 233				
<b>Experimental info</b>					
<input type="checkbox"/> Sequence conflict	95 – 96	2	Missing AA sequence <a href="#">Ref.2</a>		

Secondary structure

1. .... 25

■ Helix □ Strand ■ Turn

Details...

Sequences

Sequence	Length	Mass (Da)
<input type="checkbox"/> P00775 [UniParc]. Last modified February 1, 1994. Version 2. Checksum: 050233AFF1F64823	FASTA	259 26,776
<div><div><div>102030405060</div><div>MKHFLRALKR CSVAVATVAI AVVGLQPVTA SAAPNPVVGG TRAAQGEFPF MVRLSMGCGG</div></div><div><div>708090100110120</div><div>ALYAQDIVLT AAHCVSGSGN NTSITATGGV VDLQSSSAVK VRSTKVLQAP GYNGTGKDWA</div></div><div><div>130140150160170180</div><div>LIKLAQPINQ PTLKIATTTA YNQGTFTVAG WGANREGGSQ QRYLLKANVP FVSDAACRSA</div></div><div><div>190200210220230240</div><div>YGNELVANEE ICAGYPDTGG VDTQCQDSGG PMFRKDNADE WIQVGIVSWG YGCARPGYPG</div></div><div><div>250</div><div>VYTEVSTFAS AIASAARTL</div></div></div>		

« Hide

References

[1]	<b>"Molecular cloning and nucleotide sequence of Streptomyces griseus trypsin gene."</b> Kim J.C., Cha S.H., Jeong S.T., Oh S.K., Byun S.M. Biochem. Biophys. Res. Commun. 181:707-713(1991) [PubMed: 1755852] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA]. <u>Strain:</u> ATCC 10137 / IFO 3430 / NCIB 8232 / NCTC 6961.
[2]	<b>"Amino acid sequence of Streptomyces griseus trypsin. Cyanogen bromide fragments and complete sequence."</b> Olafson R.W., Jurasek L., Carpenter M.R., Smillie L.B. Biochemistry 14:1168-1177(1975) [PubMed: 804314] [Abstract] <u>Cited for:</u> PROTEIN SEQUENCE OF 37-259.
[3]	<b>"Refined crystal structure of Streptomyces griseus trypsin at 1.7-A resolution."</b> Read R.J., James M.N.G. J. Mol. Biol. 200:523-551(1988) [PubMed: 3135412] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
+	Additional computationally mapped references.

**Cross-references****Sequence databases**

<input type="radio"/> EMBL <input type="radio"/> GenBank <input type="radio"/> DDBJ	M64471 Genomic DNA. Translation: AAA26820.1. Sequence problems.
---	---

PIR	TRSMG. JQ1302.
-----	----------------

**3D structure databases**

<input type="radio"/> PDBe <input type="radio"/> RCSB PDB <input type="radio"/> PDBj	Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
	1OS8	X-ray	1.55	A	37-259	[»]
	1OSS	X-ray	1.93	A	37-259	[»]
	1SGT	X-ray	1.70	A	37-259	[»]
	2FMJ	X-ray	1.65	A	37-251	[»]
	3BEU	X-ray	1.05	A/B	37-259	[»]
	3I77	X-ray	2.10	A	37-259	[»]
	3I78	X-ray	3.00	A	37-259	[»]

ProteinModelPortal	P00775.
--------------------	---------

SMR	P00775. Positions 37-259.
-----	---------------------------

ModBase	Search...
---------	-----------

**Protein family/group databases**

MEROPS	S01.101.
--------	----------

**Enzyme and pathway databases**

BRENDA	3.4.21.4. 1270.
--------	-----------------

**Family and domain databases**

InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
----------	---

Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
------	--

PRINTS	PR00722. CHYMOTRYPSIN.
--------	------------------------

SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
-------	---

SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
--------	--------------------------------

PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
---------	--

ProtoNet	Search...
----------	-----------



**Entry information**

Entry name	TRYP_STRGR
Accession	Primary (citable) accession number: <b>P00775</b>
Entry history	Integrated into      July 21, 1986 UniProtKB/Swiss- Prot: Last sequence      February 1, 1994 update: Last modified:      January 11, 2011 This is version 80 of the entry and version 2 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Prokaryotic Protein Annotation Program

**Relevant documents**

PDB cross-references  
Index of Protein Data Bank (PDB) cross-references

---

Peptidase families  
Classification of peptidase families and list of entries

---

SIMILARITY comments  
Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



**P35049 (TRYP\_FUSOX) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified January 11, 2011. Version 78.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Trypsin</b> EC=3.4.21.4
Organism	<b><i>Fusarium oxysporum</i> (Panama disease fungus)</b>
Taxonomic identifier	5507 [NCBI]
Taxonomic lineage	Eukaryota › Fungi › Dikarya › Ascomycota › Pezizomycotina › Sordariomycetes › Hypocreomycetidae › Hypocreales › mitosporic Hypocreales › <i>Fusarium</i> › <i>Fusarium oxysporum</i> species complex

**Protein attributes**

Sequence length	248 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Catalytic activity	Preferential cleavage: Arg- -Xaa, Lys- -Xaa.
Subcellular location	Secreted.
Sequence similarities	Belongs to the peptidase S1 family. Contains 1 peptidase S1 domain.

**Ontologies****Keywords**












Cellular component	Secreted
Domain	Signal
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Zymogen
Technical term	3D-structure

**Gene Ontology (GO)**

Biological process	proteolysis
--------------------	-------------

Cellular component	Inferred from electronic annotation. Source: InterPro
	extracellular region Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro
Complete GO annotation...	

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
Molecule processing					
 Signal peptide	1 – 17	17	Potential		
 Propeptide	18 – 25	8	Activation peptide		PRO_0000028301
 Chain	26 – 248	223	Trypsin		PRO_0000028302
Regions					
 Domain	25 – 248	224	Peptidase S1		
Sites					
 Active site	65	1	Charge relay system		
 Active site	108	1	Charge relay system		
 Active site	204	1	Charge relay system		
 Site	198	1	Required for specificity		
Amino acid modifications					
 Disulfide bond	50 ↔ 66				
 Disulfide bond	174 ↔ 189				
 Disulfide bond	200 ↔ 225				
Secondary structure					
1	.....				248



■ Helix   ■ Strand   ■ Turn

Details...

## Sequences

Sequence		Length	Mass (Da)
	P35049 [UniParc]. FASTA	248	24,576

**Last modified February 1, 1994. Version 1.  
Checksum: 1A0EBA88C3E70294**

<u>10</u>	<u>20</u>	<u>30</u>	<u>40</u>	<u>50</u>	<u>60</u>
MVKFASVVAL	VAPLAAAAPQ	EIPNIVGGTS	ASAGDFPFIV	SISRNGGPWC	GGSLLNANTV
<u>70</u>	<u>80</u>	<u>90</u>	<u>100</u>	<u>110</u>	<u>120</u>
LTAAHCVSGY	AQSGFQIRAG	SLSRTSGGIT	SSLSSVRVHP	SYSGNNNDLA	ILKLSTSIPS
<u>130</u>	<u>140</u>	<u>150</u>	<u>160</u>	<u>170</u>	<u>180</u>
GGNIGYARLA	ASGSDPVAGS	SATVAGWGAT	SEGGSSTPVN	LLKVTVPIVS	RATCRAQYGT
<u>190</u>	<u>200</u>	<u>210</u>	<u>220</u>	<u>230</u>	<u>240</u>
SAITNQMFCA	GVSSGGKDSC	QGDSGGPIVD	SSNTLIGAVS	WGNGCARPNY	SGVYASVGAL
RSFIDTYA					

« Hide

## References

- |     |   |
|-----|---|
| [1] | <p><b>"The sequence and X-ray structure of the trypsin from <i>Fusarium oxysporum</i>."</b><br/> Rypniewski W.R., Hastrup S., Betzel C., Dauter M., Dauter Z., Papendorf G., Branner S., Wilson K.S.<br/> Protein Eng. 6:341-348(1993) [PubMed: 8332590] [Abstract]<br/> <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA], X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).</p> |
| [2] | <p><b>"Structure of inhibited trypsin from <i>Fusarium oxysporum</i> at 1.55 Å."</b><br/> Rypniewski W.R., Dambmann C., von der Osten C., Dauter M., Wilson K.S.<br/> Acta Crystallogr. D 51:73-84(1995) [PubMed: 15299338] [Abstract]<br/> <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (1.55 ANGSTROMS).</p>   |
| +   | Additional computationally mapped references.   |

## Cross-references

## Sequence databases

 EMBL  
 GenBank  
 DDBJ

**S63827 mRNA. Translation: AAB27568.1.**

**3D structure databases**

- ⊙ PDBe
- ⊙ RCSB PDB
- ⊙ PDBj

Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
1FN8	X-ray	0.81	A	25-248	[»]
1FY4	X-ray	0.81	A	25-248	[»]
1FY5	X-ray	0.81	A	25-248	[»]
1GDN	X-ray	0.81	A	25-248	[»]
1GDQ	X-ray	0.93	A	25-248	[»]
1GDU	X-ray	1.07	A	25-248	[»]
1PPZ	X-ray	1.23	A	25-248	[»]
1PQ5	X-ray	0.85	A	25-248	[»]
1PQ7	X-ray	0.80	A	25-248	[»]
1PQ8	X-ray	1.00	A	25-248	[»]
1PQA	X-ray	1.23	A	25-248	[»]
1TRY	X-ray	1.55	A	25-248	[»]
1XVM	X-ray	1.10	A	25-248	[»]
1XVO	X-ray	0.84	A	25-248	[»]
2G51	X-ray	1.84	A	25-248	[»]
2G52	X-ray	1.84	A	25-248	[»]
2VU8	X-ray	1.80	E	25-248	[»]

ProteinModelPortal

P35049.

SMR

P35049. Positions 25-248.

ModBase

Search...

**Protein family/group databases**

MEROPS

S01.103.

**Enzyme and pathway databases**

BRENDA

3.4.21.4. 15244.

**Family and domain databases**

InterPro

IPR009003. Pept\_cys/ser\_Trypsin-like.  
 IPR018114. Peptidase\_S1/S6\_AS.  
 IPR001254. Peptidase\_S1\_S6.  
 IPR001314. Peptidase\_S1A.  
 [Graphical view]

Pfam

PF00089. Trypsin. 1 hit.  
 [Graphical view]

PRINTS

PR00722. CHYMOTRYPSIN.

SMART

SM00020. Tryp\_SPc. 1 hit.  
 [Graphical view]

SUPFAM

SSF50494. Pept\_Ser\_Cys. 1 hit.

PROSITE

PS50240. TRYPSIN\_DOM. 1 hit.  
 PS00134. TRYPSIN\_HIS. 1 hit.  
 PS00135. TRYPSIN\_SER. 1 hit.  
 [Graphical view]

ProtoNet

Search...

**Entry information**

Entry name

TRYP\_FUSOX

Accession	Primary (citable) accession number: <b>P35049</b>
Entry history	Integrated into UniProtKB/Swiss-Prot: February 1, 1994 Last sequence update: February 1, 1994 Last modified: January 11, 2011 This is version 78 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Fungal Protein Annotation Program
<b>Relevant documents</b>	
PDB cross-references Index of Protein Data Bank (PDB) cross-references	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact





**Q99895 (CTRC\_HUMAN) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 98.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Chymotrypsin-C</b> EC=3.4.21.2 <i>Alternative name(s):</i> Caldecrin
Gene names	Name: <b>CTRC</b> Synonyms: CLCR
Organism	<b>Homo sapiens (Human)</b> [Complete proteome]
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo

**Protein attributes**

Sequence length	268 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Has chymotrypsin-type protease activity and hypocalcemic activity.
Catalytic activity	Preferential cleavage: Leu- -Xaa, Tyr- -Xaa, Phe- -Xaa, Met- -Xaa, Trp- -Xaa, Gln- -Xaa, Asn- -Xaa.
Tissue specificity	Pancreas.
Involvement in disease	Variations in CTRC influence susceptibility to chronic pancreatitis [MIM:167800]. Chronic pancreatitis is a persistent inflammatory disorder characterized by permanent destruction of the pancreatic parenchyma.
Sequence similarities	Belongs to the peptidase S1 family. Elastase subfamily. Contains 1 peptidase S1 domain.

**Ontologies****Keywords**

Coding sequence diversity	Polymorphism
Domain	Signal

**P17538 (CTRB1\_HUMAN) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 112.

**Names and origin**

Protein names	<p><i>Recommended name:</i>  <b>Chymotrypsinogen B</b>  EC=3.4.21.1</p> <p><u>Cleaved into the following 3 chains:</u></p> <ol style="list-style-type: none"> <li>1. <b>Chymotrypsin B chain A</b></li> <li>2. <b>Chymotrypsin B chain B</b></li> <li>3. <b>Chymotrypsin B chain C</b></li> </ol>
Gene names	<p>Name: <b>CTRB1</b>  Synonyms:CTRB</p>
Organism	<b>Homo sapiens (Human)</b> [Complete proteome]
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo

**Protein attributes**

Sequence length	263 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at transcript level.

**General annotation (Comments)**







Catalytic activity	Preferential cleavage: Tyr- -Xaa, Trp- -Xaa, Phe- -Xaa, Leu- -Xaa.
Subcellular location	Secreted › extracellular space.
Sequence similarities	<p>Belongs to the peptidase S1 family.</p> <p>Contains 1 peptidase S1 domain.</p>

**Ontologies**

<b>Keywords</b>	
Biological process	Digestion
Cellular component	Secreted
Coding sequence diversity	Polymorphism
Domain	Signal

Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Zymogen
Technical term	Complete proteome
<b>Gene Ontology (GO)</b>	
Biological process	digestion Inferred from electronic annotation. Source: UniProtKB-KW  proteolysis Non-traceable author statement. Source: UniProtKB
Cellular component	extracellular space Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	serine-type endopeptidase activity Non-traceable author statement <a href="#">Ref.1</a> . Source: UniProtKB
Complete GO annotation...	

**Sequence annotation (Features)**

	Feature key	Position(s)	Length	Description	Graphical view	Feature id
<b>Molecule processing</b>						
	Signal peptide	1 – 18	18			
	Chain	19 – 263	245	Chymotrypsinogen B		PRO_00000
	Chain	19 – 31	13	Chymotrypsin B chain A		PRO_00000
	Chain	34 – 164	131	Chymotrypsin B chain B		PRO_00000
	Chain	167 – 263	97	Chymotrypsin B chain C		PRO_00000
<b>Regions</b>						
	Domain	34 – 261	228	Peptidase S1		
<b>Sites</b>						
<input type="checkbox"/>	Active site	75	1	Charge relay system <a href="#">By similarity</a>		
<input type="checkbox"/>	Active site	120	1	Charge relay system <a href="#">By similarity</a>		
<input type="checkbox"/>	Active site	213	1	Charge relay system <a href="#">By similarity</a>		



Amino acid modifications

<input type="checkbox"/>	Disulfide bond	19 ↔ 140	By similarity		
<input type="checkbox"/>	Disulfide bond	60 ↔ 76	By similarity		
<input type="checkbox"/>	Disulfide bond	154 ↔ 219	By similarity		
<input type="checkbox"/>	Disulfide bond	186 ↔ 200	By similarity		
<input type="checkbox"/>	Disulfide bond	209 ↔ 238	By similarity		

Natural variations

<input type="checkbox"/>	Natural variant	222	1	D → H. [dbSNP:rs8061550]		VAR_05715
<input type="checkbox"/>	Natural variant	250	1	T → A. [dbSNP:rs4737]		VAR_01456

Sequences

Sequence	Length	Mass (Da)
<input checked="" type="checkbox"/> P17538 [UniParc]. Last modified August 1, 1990. Version 1. Checksum: 4C1C055A490B8701	FASTA	263 27,870
<div><div><div>102030405060</div><div>MAFLWLLSCW ALLGTTFGCG VPAIHPVLSG LSRIVNGEDA VPGSWPWQVS LQDKTGFHFC</div></div><div><div>708090100110120</div><div>GGSLISEDWV VTAAHCGVRT SDVVVAGEFD QGSDEENIQV LKIAKVFKNP KFSILTVNND</div></div><div><div>130140150160170180</div><div>ITLLKLATPA RFSQTVSAVC LPSADDDFPA GTLCATTGWG KTKYNANKTP DKLQQAALPL</div></div><div><div>190200210220230240</div><div>LSNAECKKSW GRRITDVMIC AGASGVSSCM GDSGGPLVCQ KDGAWTLVGI VSWGSDTCST</div></div><div><div>250260</div><div>SSPGVYARVT KLIPWVQKIL AAN</div></div></div>		

« Hide

References

	« Hide 'large scale' references
[1]	"Molecular cloning and nucleotide sequence of human pancreatic prechymotrypsinogen cDNA." Tomita N., Izumoto Y., Horii A., Doi S., Yokouchi H., Ogawa M., Mori T., Matsubara K. Biochem. Biophys. Res. Commun. 158:569-575(1989) [PubMed: 2917002] [Abstract] Cited for: NUCLEOTIDE SEQUENCE [MRNA]. Tissue: Pancreas.
[2]	"Cloning of human full-length CDSs in BD Creator(TM) system donor vector." Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S., Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y., Phelan M., Farmer A. Submitted (OCT-2004) to the EMBL/GenBank/DDBJ databases Cited for: NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

- [3] "The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC)."  
 The MGC Project Team  
 Genome Res. 14:2121-2127(2004) [PubMed: 15489334] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
Tissue: Pancreas.

+ Additional computationally mapped references.

## Web resources

Wikipedia  
 Chymotrypsin entry

## Cross-references

### Sequence databases

⊙ EMBL	M24400 mRNA. Translation: AAA52128.1.
⊙ GenBank	BT007356 mRNA. Translation: AAP36020.1.
⊙ DDBJ	BC005385 mRNA. Translation: AAH05385.1.

IPI	IPI00015133.
-----	--------------

PIR	A31299.
-----	---------

UniGene	Hs.610926.
---------	------------

### 3D structure databases

ProteinModelPortal	P17538.
--------------------	---------

SMR	P17538. Positions 19-263.
-----	---------------------------

ModBase	Search...
---------	-----------

### Protein-protein interaction databases

STRING	P17538.
--------	---------

### PTM databases

PhosphoSite	P17538.
-------------	---------

### Proteomic databases

PRIDE	P17538.
-------	---------

### Genome annotation databases

Ensembl	ENST00000361017; ENSP00000354294; ENSG00000168925.
---------	--

### Organism-specific databases

GeneCards	GC16P061006.
-----------	--------------

H-InvDB	HIX0013242.
---------	-------------

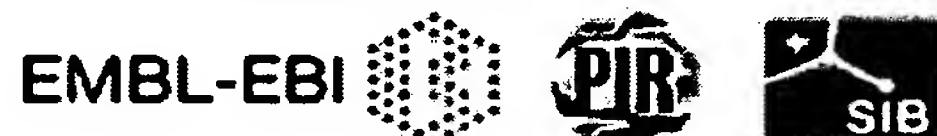
HGNC	HGNC:2521. CTRB1.
------	-------------------

MIM	118890. gene.
neXtProt	NX_P17538.
GenAtlas	Search...
<b>Phylogenomic databases</b>	
eggNOG	prNOG07094.
GeneTree	ENSGT00550000074117.
HOGENOM	HBG755338.
HOVERGEN	HBG013304.
InParanoid	P17538.
OrthoDB	EOG4VMFG1.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.21.1. 247.
<b>Gene expression databases</b>	
ArrayExpress	P17538.
CleanEx	HS_CTRB1.
Genevestigator	P17538.
GermOnline	ENSG00000168928. Homo sapiens.
<b>Family and domain databases</b>	
InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
SOURCE	Search...
<b>Entry information</b>	
Entry name	CTRB1_HUMAN
Accession	Primary (citable) accession number: <b>P17538</b>



Entry history	Integrated into    August 1, 1990 UniProtKB/Swiss- Prot: Last sequence    August 1, 1990 update: Last modified:    April 5, 2011 This is version 112 of the entry and version 1. of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
Disclaimer	Any medical or genetic information present in this entry is provided for research, educational and informational purposes only. It is not in any way intended to be used as a substitute for professional medical advice, diagnosis, treatment or care.
<b>Relevant documents</b>	
Human chromosome 16 Human chromosome 16: entries, gene names and cross-references to MIM	
Human entries with polymorphisms or disease mutations List of human entries with polymorphisms or disease mutations	
Human polymorphisms and disease mutations Index of human polymorphisms and disease mutations	
MIM cross-references Online Mendelian Inheritance in Man (MIM) cross-references in UniProtKB/Swiss-Prot	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



## Protein

Translations of Life

Display Settings: GenPept

## preprochymotrypsinogen (EC 3.4.21.1) [Homo sapiens]

GenBank: AAA52128.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAA52128 263 aa linear PRI 01-NOV-1994  
DEFINITION preprochymotrypsinogen (EC 3.4.21.1) [Homo sapiens].  
ACCESSION AAA52128  
VERSION AAA52128.1 GI:181190  
DBSOURCE locus HUMCTRP accession M24400.1  
KEYWORDS .  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 263)  
AUTHORS Tomita,N., Izumoto,Y., Horii,A., Doi,S., Yokouchi,H., Ogawa,M.,  
Mori,T. and Matsubara,K.  
TITLE Molecular cloning and nucleotide sequence of human pancreatic  
prechymotrypsinogen cDNA  
JOURNAL Biochem. Biophys. Res. Commun. 158 (2), 569-575 (1989)  
PUBMED 2917002  
COMMENT Draft entry and printed copy of sequence for [1] kindly submitted  
by N.Tomita, 07-AUG-1989.  
Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..263  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="16q23-q24.1"  
Protein 1..263  
/name="preprochymotrypsinogen (EC 3.4.21.1)"  
sig\_peptide 1..18  
/note="chymotrypsinogen signal peptide"  
mat\_peptide 34..263  
/product="chymotrypsinogen"  
Region 34..259  
/region\_name="Tryp\_SpC"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cd00190"  
/db\_xref="CDD:29152"  
Site 34  
/site\_type="cleavage"  
/db\_xref="CDD:29152"  
Site order(75,120,213)  
/site\_type="active"  
/db\_xref="CDD:29152"  
Site order(207,232,234)  
/site\_type="other"  
/note="substrate binding sites"  
/db\_xref="CDD:29152"  
CDS 1..263  
/gene="CTRB"  
/coded\_by="M24400.1:17..808"  
/db\_xref="GDB:G00-119-820"

ORIGIN  
1 maflwllscw allgttfqcg vpaihpvlsg lsriuvgeda vpgswpwqvs lqdktgfhfc  
61 ggslisedwv vtaahcgvrt sdvvvagefd qgsdeenigv lkiakvfknk kfsiltvnnd  
121 itllklatpa rfsqtvavc lpsadddfp gtlcattgw ktkynanktp dklqqaalpl  
181 lsnaeckksw grritdvmic agasgvsscm gdsggplvcq kdgawtlvgi vswgsdtcst  
241 sspgvyarvt klipwqkil aan  
//

## Protein

Translations of Life

Display Settings: GenPept

## chymotrypsinogen B1 [Homo sapiens]

GenBank: AAP36020.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAP36020 263 aa linear PRI 13-MAY-2003  
DEFINITION chymotrypsinogen B1 [Homo sapiens].  
ACCESSION AAP36020  
VERSION AAP36020.1 GI:30583551  
DBSOURCE accession BT007356.1  
KEYWORDS .  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 263)  
AUTHORS Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,  
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,  
Phelan,M. and Farmer,A.  
TITLE Cloning of human full-length CDSs in BD Creator(TM) System Donor  
vector  
JOURNAL Unpublished  
REFERENCE 2 (residues 1 to 263)  
AUTHORS Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,  
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,  
Phelan,M. and Farmer,A.  
TITLE Direct Submission  
JOURNAL Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow  
Circle, Palo Alto, CA 94303, USA  
COMMENT Method: conceptual translation supplied by author.  
FEATURES  
source Location/Qualifiers  
1..263  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="GH00216X1.0"  
/clone\_lib="BD Creator(TM) CDS Library derived from MGC  
collection"  
/lab\_host="DH5alpha T1 resistant"  
/note="Vector: pDNR-Dual"  
Protein 1..263  
/product="chymotrypsinogen B1"  
Region 34..259  
/region\_name="Tryp\_SPC"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cd00190"  
/db\_xref="CDD:29152"  
Site 34  
/site\_type="cleavage"  
/db\_xref="CDD:29152"  
Site order(75,120,213)  
/site\_type="active"  
/db\_xref="CDD:29152"  
Site order(207,232,234)  
/site\_type="other"  
/note="substrate binding sites"  
/db\_xref="CDD:29152"  
CDS 1..263  
/coded\_by="BT007356.1:1..792"  
ORIGIN  
1 mafllwllscw allgttfgcg vpaihpvlsg lsrivngeda vpgswpwqvs lqdktgfhfc  
61 ggslisedwv vtaahcgvrt sdvvvagefd qgsdeenigv lkiakvfknk kfsiltvnnd  
121 itllklatpa rfsqtvsavc lpsadddfp a gtlcattgwg ktkynanktp dklqqaalpl  
181 lsnaecksw grritdvmic agasgvsscm gdsggplvcq kdgawtlvgi vswgsdtcst  
241 sspgvyarvt klipwvqkil aan  
//



**P07338 (CTRB1\_RAT) ★ Reviewed, UniProtKB/Swiss-Prot**




Last modified April 5, 2011. Version 97.

Names and origin	
Protein names	<i>Recommended name:</i> <b>Chymotrypsinogen B</b> EC=3.4.21.1 <i>Cleaved into the following 3 chains:</i> 1. <b>Chymotrypsin B chain A</b> 2. <b>Chymotrypsin B chain B</b> 3. <b>Chymotrypsin B chain C</b>
Gene names	Name: <b>Ctrb1</b> Synonyms: Ctrb
Organism	<b>Rattus norvegicus (Rat)</b>
Taxonomic identifier	10116 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires › Rodentia › Sciurognathi › Muroidea › Muridae › Murinae › Rattus
Protein attributes	
Sequence length	263 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.
General annotation (Comments)	
Catalytic activity	Preferential cleavage: Tyr- -Xaa, Trp- -Xaa, Phe- -Xaa, Leu- -Xaa.
Subcellular location	Secreted › extracellular space.
Sequence similarities	Belongs to the peptidase S1 family. Contains 1 peptidase S1 domain.
Ontologies	
Keywords	
Biological process	Digestion
Cellular component	Secreted
Domain	Signal

Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Zymogen
Technical term	3D-structure
Gene Ontology (GO) Biological process	digestion Inferred from electronic annotation. Source: UniProtKB-KW  positive regulation of apoptosis Inferred from mutant phenotype. Source: RGD  protein catabolic process Inferred from direct assay. Source: RGD  response to cytokine stimulus Inferred from direct assay. Source: RGD  response to food Inferred from expression pattern. Source: RGD  response to nutrient Inferred from expression pattern. Source: RGD  response to peptide hormone stimulus Inferred from expression pattern. Source: RGD  response to toxin Inferred from expression pattern. Source: RGD
Cellular component	extracellular space Inferred from electronic annotation. Source: UniProtKB-SubCell  lysosome Inferred from direct assay. Source: RGD
Molecular function	protein binding Inferred from physical interaction. Source: RGD  serine-type endopeptidase activity Inferred from direct assay. Source: RGD

Complete GO annotation...

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature ident
Molecule processing					
	Signal peptide	1 – 18	18		
	Chain	19 – 263	245 Chymotrypsinogen B		PRO_000002
	Chain	19 – 31	13 Chymotrypsin B chain A		PRO_000002

	Chain	34 – 164	131	Chymotrypsin B chain B		PRO_000002
	Chain	167 – 263	97	Chymotrypsin B chain C		PRO_000002

Regions

	Domain	34 – 261	228	Peptidase S1		
--	--------	----------	-----	--------------	--	--

Sites

<input type="checkbox"/>	Active site	75	1	Charge relay system By similarity		
<input type="checkbox"/>	Active site	120	1	Charge relay system By similarity		
<input type="checkbox"/>	Active site	213	1	Charge relay system By similarity		

Amino acid modifications

<input type="checkbox"/>	Disulfide bond	19 ↔ 140	By similarity			
<input type="checkbox"/>	Disulfide bond	60 ↔ 76				
<input type="checkbox"/>	Disulfide bond	154 ↔ 219				
<input type="checkbox"/>	Disulfide bond	186 ↔ 200	By similarity			
<input type="checkbox"/>	Disulfide bond	209 ↔ 238	By similarity			

Secondary structure

1 . . . . .

■ Helix

■ Strand

■ Turn

Details...

Sequences

Sequence	Length	Mass (Da)
P07338 [UniParc]. Last modified April 1, 1988. Version 1. Checksum: ACAFDBACF8C4DA6D	FASTA	263 27,849
<div><div>102030405060</div><div>MAFLWLVS</div><div>CFALVGATFGCGVPTIQPVL</div><div>TLGLSRIVNGEDAIPGSWPWQVS</div><div>LQDKTGFHFC</div></div> <div><div>708090100110120</div><div>GGSLISEDWV</div><div>VTAAHCGVKTS</div><div>SDVVVAGEFDQGSDEENIQV</div><div>LKIAQVFKNPKFNMFTVRND</div></div> <div><div>130140150160170180</div><div>ITLLKLATPA</div><div>QFSETVSAVC</div><div>LPNVDDDFPPGTVCA</div><div>TGWGKTKYNALKTPEKLQQAALPI</div></div> <div><div>190200210220230240</div><div></div><div></div><div></div><div></div><div></div></div>		



VSEADCKKSW GSKITDVMTC AGASGVSSCM GDSGGPLVCQ KDGWVTLGI VSWGSGVCST

250 260

STPAVYSRV<sup>T</sup> ALMPWVQ<sup>L</sup>IL EAN

« Hide

References

[1]	"Isolation and sequence of a rat chymotrypsin B gene." Bell G.I., Quinto C., Quiroga M., Valenzuela P., Craik C.S., Rutter W.J. J. Biol. Chem. 259:14265-14270(1984) [PubMed: 6209274] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA].
+	Additional computationally mapped references.

Cross-references

Sequence databases						
<div><div><div>🔍 EMBL</div><div>🔍 GenBank</div><div>🔍 DDBJ</div></div></div>	K02298 Genomic DNA. Translation: AAA98732.1.					
IPI	IPI00206309.					
PIR	KYRTB. A22658.					
RefSeq	NP_036668.1. NM_012536.1.					
UniGene	Rn.105845.					
3D structure databases						
<div><div><div>🔍 PDBe</div><div>🔍 RCSB PDB</div><div>🔍 PDBj</div></div></div>	Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
	1KDQ	X-ray	2.55	A	34-164	[»]
				B	165-263	[»]
	2JET	X-ray	2.20	A	19-28	[»]
				B	37-164	[»]
				C	165-263	[»]
ProteinModelPortal	P07338.					
SMR	P07338. Positions 19-263.					
ModBase	Search...					
Protein-protein interaction databases						
STRING	P07338.					
Protein family/group databases						
MEROPS	S01.152.					
PTM databases						
PhosphoSite	P07338.					

**Genome annotation databases**

Ensembl	ENSRNOT00000026017; ENSRNOP00000026017; ENSRNOG00000019068.
---------	---

GeneID	24291.
--------	--------

KEGG	rno:24291.
------	------------

**Organism-specific databases**

CTD	24291.
-----	--------

RGD	2444. Ctrb1.
-----	--------------

**Phylogenomic databases**

eggNOG	roNOG11118.
--------	-------------

GeneTree	ENSGT00550000074117.
----------	----------------------

HOVERGEN	HBG013304.
----------	------------

InParanoid	P07338.
------------	---------

OrthoDB	EOG4VMFG1.
---------	------------

**Enzyme and pathway databases**

BRENDA	3.4.21.1. 248.
--------	----------------

**Gene expression databases**

ArrayExpress	P07338.
--------------	---------

Genevestigator	P07338.
----------------	---------

GermOnline	ENSRNOG00000019068. Rattus norvegicus.
------------	--

**Family and domain databases**

InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
----------	---

Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
------	--

PRINTS	PR00722. CHYMOTRYPSIN.
--------	------------------------

SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
-------	---

SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
--------	--------------------------------

PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
---------	--

ProtoNet	Search...
----------	-----------

**Other Resources**

NextBio	602890.
---------	---------

**Entry information**

Entry name	CTRB1_RAT
Accession	Primary (citable) accession number: <b>P07338</b>
Entry history	Integrated into April 1, 1988 UniProtKB/Swiss-Prot: Last sequence April 1, 1988 update: Last modified: April 5, 2011 This is version 97 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program

**Relevant documents**

PDB cross-references  
Index of Protein Data Bank (PDB) cross-references

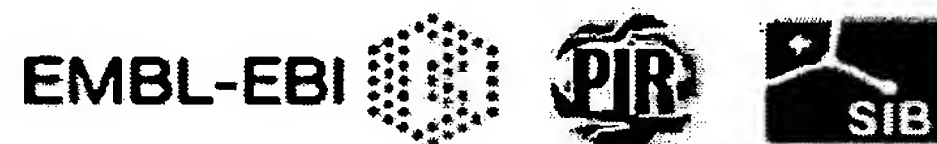
---

Peptidase families  
Classification of peptidase families and list of entries

---

SIMILARITY comments  
Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact





**P80646 (CTRB\_GADMO) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified January 11, 2011. Version 61.

Names and origin	
Protein names	<i>Recommended name:</i> <b>Chymotrypsin B</b> EC=3.4.21.1 <u>Cleaved into the following 2 chains:</u> 1. <b>Chymotrypsin B chain A</b> 2. <b>Chymotrypsin B chain B</b>
Organism	<b>Gadus morhua (Atlantic cod)</b>
Taxonomic identifier	8049 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Actinopterygii › Neopterygii › Teleostei › Euteleostei › Neoteleostei › Acanthomorpha › Paracanthopterygii › Gadiformes › Gadidae › Gadus
Protein attributes	
Sequence length	245 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.
General annotation (Comments)	
Catalytic activity	Preferential cleavage: Tyr- -Xaa, Trp- -Xaa, Phe- -Xaa, Leu- -Xaa.
Subcellular location	Secreted › extracellular space.
Sequence similarities	Belongs to the peptidase S1 family. Contains 1 peptidase S1 domain.
Ontologies	
Keywords	
Biological process	Digestion
Cellular component	Secreted
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Zymogen

Technical term		Direct protein sequencing				
Gene Ontology (GO)						
Biological process		digestion Inferred from electronic annotation. Source: UniProtKB-KW  proteolysis Inferred from electronic annotation. Source: InterPro				
Cellular component		extracellular space Inferred from electronic annotation. Source: UniProtKB-SubCell				
Molecular function		serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro				
Complete GO annotation...						
Sequence annotation (Features)						
	Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
Molecule processing						
<input checked="" type="checkbox"/>	Chain	1 – 13	13	Chymotrypsin B chain A		PRO_0000027
<input type="checkbox"/>	Propeptide	14 – 15	2			PRO_0000027
<input checked="" type="checkbox"/>	Chain	16 – 245	230	Chymotrypsin B chain B		PRO_0000027
Regions						
<input checked="" type="checkbox"/>	Domain	16 – 243	228	Peptidase S1		
Sites						
<input type="checkbox"/>	Active site	57	1	Charge relay system By similarity		
<input type="checkbox"/>	Active site	101	1	Charge relay system By similarity		
<input type="checkbox"/>	Active site	195	1	Charge relay system By similarity		
Amino acid modifications						
<input type="checkbox"/>	Disulfide bond	1 ↔ 121		By similarity		
<input type="checkbox"/>	Disulfide bond	42 ↔ 58		By similarity		
<input type="checkbox"/>	Disulfide bond	135 ↔ 201		By similarity		
<input type="checkbox"/>	Disulfide bond	167 ↔ 182		By similarity		

<input type="checkbox"/>	Disulfide bond	191 ↔ 220		By similarity		
Experimental info						
<input type="checkbox"/>	Sequence conflict	9 – 11	3	QVT → VIS AA sequence Ref.2		
<input type="checkbox"/>	Sequence conflict	26	1	S → T AA sequence Ref.2		
<input type="checkbox"/>	Sequence conflict	28 – 29	2	PW → Y AA sequence Ref.2		

Sequences

Sequence	Length	Mass (Da)
<input checked="" type="checkbox"/> P80646 [UniParc]. Last modified October 1, 1996. Version 1. Checksum: 74FE0D425517AB02	FASTA 245	26,260
<div><div><div>102030405060</div><div>CGSPAIQPQVTGYARIVNGEEAVPHSWPWQVSLQQSNGFHFCCGSLINENWVVTAAHCNV</div></div><div><div>708090100110120</div><div>RTYHRVIVGEHDKASDENIQILKPSMVFTHPKWDSRTINNDISLIKLASPAVLGTVNSPV</div></div><div><div>130140150160170180</div><div>CLGESSDVFAPGMKCVTSGWGLTRYNAPGTPNKLQQAALPLMSNEECSQTWGNNMISDVM</div></div><div><div>190200210220230240</div><div>ICAGAAGATS CMGDSGGPLVCQKDNVWTLVGIVSWGSSRC SVTTPAVYARVTELRGWVDQ</div></div><div>ILAAAN</div></div>		

« Hide

References

[1]	<b>"Structure of chymotrypsin variant B from Atlantic cod, Gadus morhua."</b> Leth-Larsen R., Asgeirsson B., Thorolfsson M., Noerregaard-Madsen M., Hoejrup P. Biochim. Biophys. Acta 1297:49-56(1996) [PubMed: 8841380] [Abstract] <u>Cited for:</u> PROTEIN SEQUENCE. <u>Tissue:</u> Pyloric caecum.
[2]	<b>"Structural and kinetic properties of chymotrypsin from Atlantic cod (Gadus morhua). Comparison with bovine chymotrypsin."</b> Asgeirsson B., Bjarnason J.B. Comp. Biochem. Physiol. 99B:327-335(1991) [PubMed: 1764912] [Abstract] <u>Cited for:</u> PROTEIN SEQUENCE OF 1-12 AND 16-31. <u>Tissue:</u> Pyloric caecum.



**Cross-references****3D structure databases**

ProteinModelPortal	P80646.
SMR	P80646. Positions 1-245.
ModBase	Search...

**Protein family/group databases**

MEROPS	S01.437.
--------	----------

**Phylogenomic databases**

HOVERGEN	HBG013304.
----------	------------

**Enzyme and pathway databases**

BRENDA	3.4.21.1. 39336.
--------	------------------

**Family and domain databases**

InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...

**Entry information**

Entry name	CTRB_GADMO
Accession	Primary (citable) accession number: <b>P80646</b>
Entry history	Integrated into UniProtKB/Swiss-Prot: October 1, 1996 Last sequence update: October 1, 1996 Last modified: January 11, 2011 This is version 61 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)

Annotation program

Chordata Protein Annotation Program

**Relevant documents**

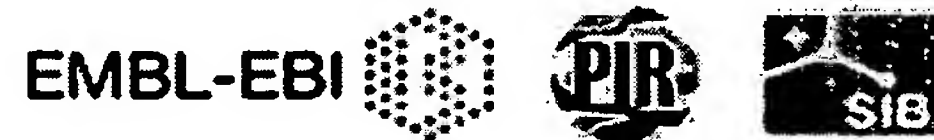
Peptidase families

Classification of peptidase families and list of entries

SIMILARITY comments

Index of protein domains and families

© 2002–2011 UniProt Consortium | License &amp; Disclaimer | Contact



**P00768 (CTR2\_VESOR) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified January 11, 2011. Version 71.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Chymotrypsin-2</b> EC=3.4.21.1 <i>Alternative name(s):</i> Chymotrypsin II
Organism	<b>Vespa orientalis (Oriental hornet)</b>
Taxonomic identifier	7447 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Arthropoda › Hexapoda › Insecta › Pterygota › Neoptera › Endopterygota › Hymenoptera › Apocrita › Aculeata › Vespoidea › Vespidae › Vespinae › Vespa

**Protein attributes**

Sequence length	216 AA.
Sequence status	Complete.
Protein existence	Evidence at protein level.










**General annotation (Comments)**

Catalytic activity	Preferential cleavage: Tyr- -Xaa, Trp- -Xaa, Phe- -Xaa, Leu- -Xaa.
Subcellular location	Secreted › extracellular space.
Sequence similarities	Belongs to the peptidase S1 family. Contains 1 peptidase S1 domain.

**Ontologies**

<b>Keywords</b>	
Cellular component	Secreted
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond
Technical term	Direct protein sequencing
<b>Gene Ontology (GO)</b>	
Biological process	proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	extracellular space



		Inferred from electronic annotation. Source: UniProtKB-SubCell				
Molecular function		serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro				
Complete GO annotation...						
Sequence annotation (Features)						
	Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
Molecule processing						
	Chain	1 – 216	216	Chymotrypsin -2		PRO_0000088677
Regions						
	Domain	1 – 216	216	Peptidase S1		
Sites						
<input type="checkbox"/>	Active site	39	1	Charge relay system <a href="#">Ref.2</a>		
<input type="checkbox"/>	Active site	82	1	Charge relay system <a href="#">Ref.2</a>		
<input type="checkbox"/>	Active site	173	1	Charge relay system <a href="#">Ref.2</a>		
Amino acid modifications						
<input type="checkbox"/>	Disulfide bond	25 ↔ 40		<a href="#">Ref.1</a>		
<input type="checkbox"/>	Disulfide bond	146 ↔ 159		<a href="#">Ref.1</a>		
<input type="checkbox"/>	Disulfide bond	169 ↔ 193		<a href="#">Ref.1</a>		
Sequences						
	Sequence		Length	Mass (Da)		
	P00768 [UniParc].	FASTA	216	23,471		
Last modified July 21, 1986. Version 1. Checksum: F235BF992AEFEDE1						
<div><div><div>102030405060</div><div>IVGGTNAPRGKYPYQVSLRA PKHFCGGSIS KRYVLTAAHC LVGKSEHQVT VGSVLLNKEE</div></div><div><div>708090100110120</div><div>AVYNAKELIV NKNYNSIRLI NDIGLIRVSK DISFTQLVQP VKLPVSNTIK AGDPVVLGTW</div></div><div><div>130140150160170180</div><div>GRIYVNGPIP NNLQQITLSI VNQQTCKSKH WGLTDSQICT FTKRGEGACH GDSGGPLVAN</div></div></div>						

190 200 210  
GVQIGIVSYG HPCAIGSPNV FTRVYSFLDW IQKNQL

« Hide

## References

- [1] **"Amino acid sequence of an insect chymotrypsin from the larvae of the hornet, *Vespa orientalis*."**  
Jany K.-D., Bekelar K., Pfeleiderer G., Ishay J.  
Biochem. Biophys. Res. Commun. 110:1-7(1983) [PubMed: 6340663] [Abstract]  
Cited for: PROTEIN SEQUENCE, DISULFIDE BONDS.
- [2] **"The amino acid sequences around the reactive serine and histidine residues of the chymotrypsin-like protease from the hornet, *Vespa orientalis*."**  
Jany K.-D., Bekelar K., Ishay J.  
Biochim. Biophys. Acta 668:197-200(1981) [PubMed: 6786354] [Abstract]  
Cited for: ACTIVE SITE.

## Cross-references

### Sequence databases

PIR | KYVH2O. A00954.

### 3D structure databases

ProteinModelPortal | P00768.

SMR | P00768. Positions 1-214.

ModBase | Search...

### Protein family/group databases

MEROPS | S01.438.

### Enzyme and pathway databases

BRENDA | 3.4.21.1. 290116.

### Family and domain databases

InterPro | IPR009003. Pept\_cys/ser\_Trypsin-like.  
IPR018114. Peptidase\_S1/S6\_AS.  
IPR001254. Peptidase\_S1\_S6.  
IPR001314. Peptidase\_S1A.  
[Graphical view]

Pfam | PF00089. Trypsin. 1 hit.  
[Graphical view]

PRINTS | PR00722. CHYMOTRYPSIN.

SMART | SM00020. Tryp\_SPc. 1 hit.  
[Graphical view]

SUPFAM | SSF50494. Pept\_Ser\_Cys. 1 hit.

PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Entry information</b>	
Entry name	CTR2_VESOR
Accession	Primary (citable) accession number: <b>P00768</b>
Entry history	Integrated into July 21, 1986 UniProtKB/Swiss-Prot: Last sequence July 21, 1986 update: Last modified: January 11, 2011 This is version 71 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
<b>Relevant documents</b>	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License &amp; Disclaimer | Contact





**P00769 (CTR2\_VESCR) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified January 11, 2011. Version 70.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Chymotrypsin-2</b> EC=3.4.21.1 <i>Alternative name(s):</i> Chymotrypsin II
Organism	<b><i>Vespa crabro</i> (European hornet)</b>
Taxonomic identifier	7445 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Arthropoda › Hexapoda › Insecta › Pterygota › Neoptera › Endopterygota › Hymenoptera › Apocrita › Aculeata › Vespoidea › Vespidae › Vespinae › Vespa

**Protein attributes**

Sequence length	218 AA.
Sequence status	Complete.
Protein existence	Evidence at protein level.

**General annotation (Comments)**









Catalytic activity	Preferential cleavage: Tyr- -Xaa, Trp- -Xaa, Phe- -Xaa, Leu- -Xaa.
Subcellular location	Secreted › extracellular space.
Miscellaneous	An additional Arg at the carboxyl end was found in some of the molecules.
Sequence similarities	Belongs to the peptidase S1 family. Contains 1 peptidase S1 domain.

**Ontologies**


<b>Keywords</b>	
Cellular component	Secreted
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond
Technical term	Direct protein sequencing
<b>Gene Ontology (GO)</b>	
Biological process	proteolysis Inferred from electronic annotation. Source: InterPro

Cellular component	extracellular space Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro
Complete GO annotation...	

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
Molecule processing					
 Chain	1 – 218	218	Chymotrypsin-2		PRO_0000088676
Regions					
 Domain	1 – 218	218	Peptidase S1		
Sites					
 Active site	39	1	Charge relay system		
 Active site	84	1	Charge relay system		
 Active site	175	1	Charge relay system		
Amino acid modifications					
 Disulfide bond	25 ↔ 40				
 Disulfide bond	148 ↔ 161				
 Disulfide bond	171 ↔ 195				

Sequences

Sequence	Length	Mass (Da)
 P00769 [UniParc]. Last modified July 21, 1986. Version 1. Checksum: 509AB50DE190EB39	FASTA 218	23,677
<div><div><div>102030405060</div><div>IVGGTDAPRGKYPYQVSLRA PKHFCGGSIS KRYVLTAAHCLVGKSKHQVT VHAGSVLLNK</div></div><div><div>708090100110120</div><div>EEAVYNAEEL IVNKNYNSIR LINDIGLIRV SKDISYTLQLV QPVKLPVSNT IKAGDPVVLV</div></div><div><div>130140150160170180</div><div>GWGRIYVNGP IPNNLQQITL SIVNQQTCKF KHWGLTDSQI CTFTKLGEGA CDGDSGGPLV</div></div></div>		

190 200 210  
ANGVQIGIVS YGHPCAVGSP NVFTRVYSFL DWIQKNQL

« Hide

## References

- [1] "Amino acid sequence of the chymotryptic protease II from the larvae of the nornet, *Vespa crabro*."  
Jany K.-D., Haug H.  
FEBS Lett. 158:98-102(1983)  
Cited for: PROTEIN SEQUENCE.

## Cross-references

### Sequence databases

PIR | KYVH2C. A00955.

### 3D structure databases

ProteinModelPortal | P00769.

SMR | P00769. Positions 1-216.

ModBase | Search...

### Protein family/group databases

MEROPS | S01.438.

### Enzyme and pathway databases

BRENDA | 3.4.21.1. 296866.

### Family and domain databases

InterPro | IPR009003. Pept\_cys/ser\_Trypsin-like.  
IPR018114. Peptidase\_S1/S6\_AS.  
IPR001254. Peptidase\_S1\_S6.  
IPR001314. Peptidase\_S1A.  
[Graphical view]

Pfam | PF00089. Trypsin. 1 hit.  
[Graphical view]

PRINTS | PR00722. CHYMOTRYPSIN.

SMART | SM00020. Tryp\_SPc. 1 hit.  
[Graphical view]

SUPFAM | SSF50494. Pept\_Ser\_Cys. 1 hit.

PROSITE | PS50240. TRYPSIN\_DOM. 1 hit.  
PS00134. TRYPSIN\_HIS. 1 hit.  
PS00135. TRYPSIN\_SER. 1 hit.  
[Graphical view]

ProtoNet | Search...



**Entry information**

Entry name	CTR2_VESCR
Accession	Primary (citable) accession number: <b>P00769</b>
Entry history	Integrated into      July 21, 1986 UniProtKB/Swiss- Prot: Last sequence      July 21, 1986 update: Last modified:      January 11, 2011 This is version 70 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)

**Relevant documents**

Peptidase families  
Classification of peptidase families and list of entries

SIMILARITY comments  
Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



**O97398 (CTR\_PHACE) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified January 11, 2011. Version 57.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Chymotrypsin</b> EC=3.4.21.1
Organism	<b><i>Phaedon cochleariae</i> (Mustard beetle)</b>
Taxonomic identifier	80249 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Arthropoda › Hexapoda › Insecta › Pterygota › Neoptera › Endopterygota › Coleoptera › Polyphaga › Cucujiformia › Chrysomeloidea › Chrysomelidae › Chrysomelinae › Chrysomelini › <i>Phaedon</i>

**Protein attributes**

Sequence length	276 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at transcript level.

**General annotation (Comments)**







Function	Serine protease with chymotryptic and collagenolytic activities <a href="#">By similarity</a> <a href="#">UniProtKB Q00871</a>
Catalytic activity	Preferential cleavage: Tyr- -Xaa, Trp- -Xaa, Phe- -Xaa, Leu- -Xaa. <a href="#">UniProtKB Q00871</a>
Subcellular location	Secreted › extracellular space <a href="#">By similarity</a> <a href="#">UniProtKB Q00871</a>
Tissue specificity	Expressed in larval carcasses and gut, and adult gut. <a href="#">Ref.1</a>
Developmental stage	Ovarial and mature eggs, larvae and adult. <a href="#">Ref.1</a>
Sequence similarities	Belongs to the peptidase S1 family. Contains 1 peptidase S1 domain.

**Ontologies**

<b>Keywords</b>	
Biological process	Collagen degradation Digestion
Cellular component	Secreted
Domain	Signal

Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Glycoprotein Zymogen
Gene Ontology (GO)	
Biological process	collagen catabolic process Inferred from electronic annotation. Source: UniProtKB-KW  digestion Inferred from electronic annotation. Source: UniProtKB-KW  proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	extracellular space Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro
Complete GO annotation...	

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
Molecule processing					
	Signal peptide	1 – 16	16	Potential	
	Propeptide	17 – 45	29	Activation peptide By similarity UniProtKB Q00871	PRO_000031468
	Chain	46 – 276	231	Chymotrypsin	PRO_500014732
Regions					
	Domain	46 – 272	227	Peptidase S1	
Sites					
	Active site	89	1	Charge relay system By similarity UniProtKB P00771	
	Active site	135	1	Charge relay system By similarity UniProtKB P00771	



<input type="checkbox"/>	Active site	229	1	Charge relay system	I	
				By similarity		
				UniProtKB P00771		
Amino acid modifications						
<input type="checkbox"/>	Glycosylation	144	1	N-linked (GlcNAc...)	I	
				Potential		
<input type="checkbox"/>	Glycosylation	193	1	N-linked (GlcNAc...)	I	
				Potential		
<input type="checkbox"/>	Disulfide bond	74 ↔ 90		By similarity	II	
				UniProtKB P00771		
<input type="checkbox"/>	Disulfide bond	202 ↔ 215		By similarity	II	
				UniProtKB P00771		
<input type="checkbox"/>	Disulfide bond	225 ↔ 250		By similarity	II	
				UniProtKB P00771		

Sequences

Sequence	Length	Mass (Da)
<input type="checkbox"/> O97398 [UniParc]. Last modified May 1, 1999. Version 1. Checksum: FC5FD05DB882A1DE	FASTA	276 29,868
<div><div><div>102030405060</div><div>MKVALVVLAL FGVSLAASID NIEIPPSKNI YVEPINQPEV DPSLEIVNGQ EVVPHSIPYQ</div></div><div><div>708090100110120</div><div>IFLVASAGET SWTCGGSLIT KRYVLTAHC IQGAKSVHVT LGAHNLAKHE ASKVTVNGRS</div></div><div><div>130140150160170180</div><div>WVIHEKYDST NIDNDIGVIQ LERNLTLTRS IQLARLPSLR DVGINLEGRT ATVSGWGLTN</div></div><div><div>190200210220230240</div><div>GIFQTTDVL RANNTIISNK ECNDVFKIVQ PTEVCLSIAG GRSACSGDSG GPLVIDNVQH</div></div><div><div>250260270</div><div>GIVSYGSSYC RSTPSVFTRV SSYLNWLQTH SEWRAQ</div></div></div>		

« Hide

References

[1] "Molecular cloning of cDNAs encoding a range of digestive enzymes from a phytophagous beetle, *Phaedon cochleariae*." Girard C., Jouanin L. Insect Biochem. Mol. Biol. 29:1129-1142(1999) [PubMed: 10612046] [Abstract] Cited for: NUCLEOTIDE SEQUENCE [MRNA], TISSUE SPECIFICITY, DEVELOPMENTAL STAGE. Tissue: Larval gut.

**Cross-references****Sequence databases**

⊙ EMBL	Y17904 mRNA. Translation: CAA76928.1.
⊙ GenBank	
⊙ DDBJ	

**3D structure databases**

HSSP	HSSP built from PDB template 1NN6 based on UniProtKB P23946.
ProteinModelPortal	O97398.
SMR	O97398. Positions 46-270.
ModBase	Search...

**Enzyme and pathway databases**

BRENDA	3.4.21.1. 290922.
--------	-------------------

**Family and domain databases**

InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...

**Entry information**

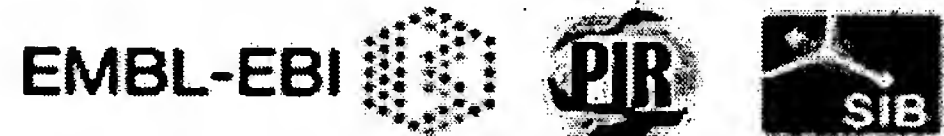
Entry name	CTR_PHACE
Accession	Primary (citable) accession number: <b>O97398</b> Secondary accession number(s): P81521
Entry history	Integrated into January 15, 2008 UniProtKB/Swiss-Prot: Last sequence May 1, 1999 update: Last modified: January 11, 2011 This is version 57 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)

**Relevant documents**

Peptidase families  
Classification of peptidase families and list of entries

SIMILARITY comments  
Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact





**P55091 (CTRC\_RAT) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified March 8, 2011. Version 83.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Chymotrypsin-C</b> EC=3.4.21.2 <i>Alternative name(s):</i> Caldecrin Serum calcium-decreasing factor
Gene names	Name: <b>Ctrc</b>
Organism	<b>Rattus norvegicus (Rat)</b>
Taxonomic identifier	10116 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires › Rodentia › Sciurognathi › Muroidea › Muridae › Murinae › Rattus

**Protein attributes**

Sequence length	268 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**





Function	Has chymotrypsin-type protease activity and hypocalcemic activity.
Catalytic activity	Preferential cleavage: Leu- -Xaa, Tyr- -Xaa, Phe- -Xaa, Met- -Xaa, Trp- -Xaa, Gln- -Xaa, Asn- -Xaa.
Tissue specificity	Pancreas.
Sequence similarities	Belongs to the peptidase S1 family. Elastase subfamily. Contains 1 peptidase S1 domain.
Caution	Was originally ( <a href="#">Ref.2</a> ) thought to be elastase IV.

**Ontologies**

<b>Keywords</b>	
Domain	Signal
Molecular function	Hydrolase Protease Serine protease

PTM	Disulfide bond Glycoprotein Zymogen
Technical term	Direct protein sequencing
<b>Gene Ontology (GO)</b>	
Biological process	proteolysis Inferred from electronic annotation. Source: InterPro
Molecular function	serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro
Complete GO annotation...	

**Sequence annotation (Features)**

	Feature key	Position(s)	Length	Description	Graphical view
<b>Molecule processing</b>					
	Signal peptide	1 – 16	16	Potential	
	Propeptide	17 – 29	13	Activation peptide	
	Chain	30 – 268	239	Chymotrypsin-C	
<b>Regions</b>					
	Domain	30 – 267	238	Peptidase S1	
<b>Sites</b>					
<input type="checkbox"/>	Active site	74	1	Charge relay system By similarity	
<input type="checkbox"/>	Active site	121	1	Charge relay system By similarity	
<input type="checkbox"/>	Active site	216	1	Charge relay system By similarity	
<b>Amino acid modifications</b>					
<input type="checkbox"/>	Glycosylation	25	1	N-linked (GlcNAc...) Potential	
<input type="checkbox"/>	Glycosylation	90	1	N-linked (GlcNAc...) Potential	
<input type="checkbox"/>	Disulfide bond	17 ↔ 141		By similarity	
<input type="checkbox"/>	Disulfide bond	59 ↔ 75		By similarity	
<input type="checkbox"/>	Disulfide bond	155 ↔ 222		By similarity	
<input type="checkbox"/>	Disulfide bond	186 ↔ 202		By similarity	
<input type="checkbox"/>	Disulfide bond	212 ↔ 243		By similarity	

## Experimental info

<input type="checkbox"/>	Sequence conflict	42	1	P → A in CAA41753. <a href="#">Ref.2</a>	
<input checked="" type="checkbox"/>	Sequence conflict	96 – 120	25	EEGSV...LFLWN → AEAPCTLRWTPSTSMRSGTD SSCGT in CAA41753. <a href="#">Ref.2</a>	

## Sequences

Sequence	Length	Mass (Da)
<input checked="" type="checkbox"/> P55091 [UniParc]. Last modified October 1, 1996. Version 1. Checksum: 33B67AF34D0F8583	FASTA 268	29,374

```
      10      20      30      40      50      60
MLGITVLAAI LACASCCGNP AFPPNLSTRV VGGEDAVPNS WPWQVSLQYL KDDTWRHTCG

      70      80      90     100     110     120
GSLITTSHVL TAAHCINKDF TYRVGLGKYN LTVEDEEGSV YAEVDTIYVH EKWNRLFLWN

     130     140     150     160     170     180
DIAIIKLAEP VELSENTIQVA CIPEEGSLLP QDYPCYVTGW GRLWTNGPIA EVLQQGLQPI

     190     200     210     220     230     240
VSHATCSRLD WWFIKVRKTM VCAGGDGVIS ACNGDSGGPL NCQAEDGSWQ VHGIVSFGSS

     250     260
SGCNVHKKPV VFTRVSAYND WINEKIQL
```

« Hide

## References

- [1] **"Molecular cloning and expression of serum calcium-decreasing factor (caldecrin)."**  
Tomomura A., Tomomura M., Fukushima T., Akiyama M., Kubota N., Kumaki K., Nishii Y.,  
Noikura T., Saheki T.  
J. Biol. Chem. 270:30315-30321(1995) [PubMed: 8530454] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA], PARTIAL PROTEIN SEQUENCE.  
Tissue: Pancreas.
- [2] **"Identification of cDNAs encoding two novel rat pancreatic serine proteases."**  
Kang J., Wiegand U., Mueller-Hill B.  
Gene 110:181-187(1992) [PubMed: 1537555] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA].  
Tissue: Pancreas.
- [3] **"Caldecrin is a novel-type serine protease expressed in pancreas, but its homologue,  
elastase IV, is an artifact during cloning derived from caldecrin gene."**  
Yoshino-Yasuda I., Kobayashi K., Akiyama M., Itoh H., Tomomura A., Saheki T.  
J. Biochem. 123:546-554(1998) [PubMed: 9538241] [Abstract]  
Cited for: CHARACTERIZATION.



**Cross-references****Sequence databases**

<input checked="" type="radio"/> EMBL <input type="radio"/> GenBank <input type="radio"/> DDBJ	S80379 mRNA. Translation: AAB35830.1. X59014 mRNA. Translation: CAA41753.1.
--	--

IPI	IPI00213599.
-----	--------------

PIR	JQ1473.
-----	---------

RefSeq	NP_001071117.1. NM_001077649.1.
--------	---------------------------------

UniGene	Rn.19214.
---------	-----------

**3D structure databases**

ProteinModelPortal	P55091.
--------------------	---------

SMR	P55091. Positions 17-268.
-----	---------------------------

ModBase	Search...
---------	-----------

**Protein-protein interaction databases**

STRING	P55091.
--------	---------

**Protein family/group databases**

MEROPS	S01.157.
--------	----------

**Proteomic databases**

PRIDE	P55091.
-------	---------

**Genome annotation databases**

Ensembl	ENSRNOT00000018393; ENSRNOP00000018393; ENSRNOG00000013745.
---------	--

GeneID	362653.
--------	---------

KEGG	rno:362653.
------	-------------

UCSC	S80379. rat.
------	--------------

**Organism-specific databases**

CTD	362653.
-----	---------

RGD	1308379. Ctrc.
-----	----------------

**Phylogenomic databases**

eggNOG	roNOG04592.
--------	-------------

GeneTree	ENSGT00550000074117.
----------	----------------------

HOVERGEN	HBG013304.
----------	------------

InParanoid	P55091.
------------	---------

OMA	WNDIAII.
-----	----------

OrthoDB	EOG4XPQGN.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.21.2. 248.
<b>Gene expression databases</b>	
ArrayExpress	P55091.
Genevestigator	P55091.
GermOnline	ENSRNOG00000013745. Rattus norvegicus.
<b>Family and domain databases</b>	
InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
NextBio	680758.
<b>Entry information</b>	
Entry name	CTRC_RAT
Accession	Primary (citable) accession number: <b>P55091</b> Secondary accession number(s): Q63188
Entry history	Integrated into      October 1, 1996 UniProtKB/Swiss- Prot: Last sequence      October 1, 1996 update: Last modified:      March 8, 2011 This is version 83 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program

**Relevant documents**

Peptidase families  
Classification of peptidase families and list of entries

SIMILARITY comments  
Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact





**P04189 (SUBT\_BACSU) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified February 8, 2011. Version 108.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Subtilisin E</b> EC=3.4.21.62
Gene names	Name: <b>aprE</b> Synonyms: <b>apr, aprA, sprE</b> Ordered Locus Names: <b>BSU10300</b>
Organism	<b>Bacillus subtilis</b> [Complete proteome] [HAMAP]
Taxonomic identifier	1423 [NCBI]
Taxonomic lineage	Bacteria › Firmicutes › Bacillales › Bacillaceae › Bacillus

**Protein attributes**

Sequence length	381 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Subtilisin is an extracellular alkaline serine protease, it catalyzes the hydrolysis of proteins and peptide amides.
Catalytic activity	Hydrolysis of proteins with broad specificity for peptide bonds, and a preference for a large uncharged residue in P1. Hydrolyzes peptide amides.
Cofactor	Binds 2 calcium ions per subunit. <a href="#">Ref.12</a>
Subcellular location	Secreted.
Domain	The propeptide functions as an intramolecular chaperone which is essential for the correct folding of the protease domain but is not required for enzymatic function of the folded protein. It is autoprocessed and degraded after completion of the folding process. <a href="#">Ref.11</a>
Miscellaneous	Secretion of subtilisin is associated with onset of sporulation, and many mutations which block sporulation at early stages affect expression levels of subtilisin. However, subtilisin is not necessary for normal sporulation.
Sequence similarities	Belongs to the peptidase S8 family.

Ontologies

Keywords




Biological process	Sporulation
Cellular component	Secreted
Domain	Signal
Ligand	Calcium Metal-binding
Molecular function	Hydrolase Protease Serine protease
PTM	Autocatalytic cleavage Zymogen
Technical term	3D-structure Complete proteome Direct protein sequencing

Gene Ontology (GO)

Biological process	negative regulation of catalytic activity Inferred from electronic annotation. Source: InterPro  proteolysis Inferred from electronic annotation. Source: InterPro  sporulation resulting in formation of a cellular spore Inferred from electronic annotation. Source: UniProtKB-KW
Cellular component	extracellular region Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	identical protein binding Inferred from electronic annotation. Source: InterPro  metal ion binding Inferred from electronic annotation. Source: UniProtKB-KW  serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro

Complete GO annotation...

Sequence annotation (Features)

	Feature key	Position (s)	Length	Description	Graphical view	Feature identifier
Molecule processing						
	Signal peptide	1 – 29	29	<a href="#">Ref.10</a>		
	Propeptide	30 – 106	77			PRO_0000027
	Chain	107 – 381	275	Subtilisin E		PRO_0000027

Sites					
<input type="checkbox"/>	Active site	138	1	Charge relay system	I
<input type="checkbox"/>	Active site	170	1	Charge relay system	I
<input type="checkbox"/>	Active site	327	1	Charge relay system	I
<input type="checkbox"/>	Metal binding	108	1	Calcium 1	I
<input type="checkbox"/>	Metal binding	147	1	Calcium 1	I
<input type="checkbox"/>	Metal binding	181	1	Calcium 1; via carbonyl oxygen	I
<input type="checkbox"/>	Metal binding	183	1	Calcium 1	I
<input type="checkbox"/>	Metal binding	185	1	Calcium 1; via carbonyl oxygen	I
<input type="checkbox"/>	Metal binding	187	1	Calcium 1; via carbonyl oxygen	I
<input type="checkbox"/>	Metal binding	275	1	Calcium 2; via carbonyl oxygen	I
<input type="checkbox"/>	Metal binding	277	1	Calcium 2; via carbonyl oxygen	I
<input type="checkbox"/>	Metal binding	280	1	Calcium 2	I
<input type="checkbox"/>	Metal binding	303	1	Calcium 2	I
Experimental info					
<input type="checkbox"/>	Mutagenesis	138	1	D → N: Prevents intramolecular processing to an active enzyme. Ref.11	I
<input type="checkbox"/>	Sequence conflict	27	1	A → V in CAA74536. Ref.2	I
<input type="checkbox"/>	Sequence conflict	191	1	A → S in AAA22742. Ref.1	I
<input type="checkbox"/>	Sequence conflict	191	1	A → S in CAA74536. Ref.2	I
<input type="checkbox"/>	Sequence conflict	191	1	A → S Ref.7	I



Secondary structure

1 ..... 3

■ Helix ■ Strand ■ Turn


Details...

Sequences

Sequence	Length	Mass (Da)
<div><div><div><div></div><div>P04189 [UniParc].</div></div><div>FASTA</div></div><div>381</div><div>39,479</div></div> <div>Last modified June 16, 2009. Version 3. Checksum: 5C8A4B0E42FCE83D</div> <div><div><div><div>102030405060</div><div>MRSKKLWISL LFALTTLIFTM AFSNMSAQAA GKSSTEKKYI VGFKQTMSAM SSAKKKDVIS</div></div><div><div>708090100110120</div><div>EKGGKVQKQF KYVNAAAATL DEKAVKELKK DPSVAYVEED HIAHEYAQSV PYGISQIKAP</div></div><div><div>130140150160170180</div><div>ALHSQGYTGS NVKVAVIDSG IDSSHPDLNV RGGASFVPSE TNPYQDGSSH GTHVAGTIAA</div></div><div><div>190200210220230240</div><div>LNNSIGVLGV APSASLYAVK VLDSTGSGQY SWIINGIEWA ISNNMDVINM SLGGPTGSTA</div></div><div><div>250260270280290300</div><div>LKTVVDKAVS SGIVVAAAAG NEGSSGSTST VGYPAKYPST IAVGAVNSSN QRASFSSAGS</div></div><div><div>310320330340350360</div><div>ELDVMAPGVS IQSTLPGGTY GAYNGTSMAT PHVAGAAALI LSKHPTWTNA QVRDRLESTA</div></div><div><div>370380</div><div>TYLGNSFYYG KGLINVQAAA Q</div></div></div></div> <div>« Hide</div>		

References

	« Hide 'large scale' references
[1]	<p><b>"Replacement of the Bacillus subtilis subtilisin structural gene with an In vitro-derived deletion mutation."</b> Stahl M.L., Ferrari E. J. Bacteriol. 158:411-418(1984) [PubMed: 6427178] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA]. <u>Strain:</u> 168.</p>
[2]	<p><b>"The 172 kb prkA-addAB region from 83 degrees to 97 degrees of the Bacillus subtilis chromosome contains several dysfunctional genes, the glyB marker, many genes encoding transporter proteins, and the ubiquitous hit gene."</b> Noback M.A., Holsappel S., Kiewiet R., Terpstra P., Wambutt R., Wedler H., Venema G., Bron S. Microbiology 144:859-875(1998) [PubMed: 9579061] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA]. <u>Strain:</u> 168.</p>
[3]	<p><b>"The complete genome sequence of the Gram-positive bacterium Bacillus subtilis."</b> Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Borriss R., Boursier L., Brans A., Braun M., Brignell</p>

	<p>S.C., Bron S., Brouillet S., Bruschi C.V.  Danchin A.  Nature 390:249-256(1997) [PubMed: 9384377] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  <u>Strain:</u> 168.</p>
[4]	<p><b>"From a consortium sequence to a unified sequence: the <i>Bacillus subtilis</i> 168 reference genome a decade later."</b>  Barbe V., Cruveiller S., Kunst F., Lenoble P., Meurice G., Sekowska A., Vallenet D., Wang T., Moszer I., Medigue C., Danchin A.  Microbiology 155:1758-1775(2009) [PubMed: 19383706] [Abstract]  <u>Cited for:</u> SEQUENCE REVISION TO 27 AND 191.</p>
[5]	<p><b>"The subtilisin E gene of <i>Bacillus subtilis</i> is transcribed from a sigma 37 promoter in vivo."</b>  Wong S.L., Price C.W., Goldfarb D.S., Doi R.H.  Proc. Natl. Acad. Sci. U.S.A. 81:1184-1188(1984) [PubMed: 6322190] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-156.  <u>Strain:</u> 168 / PY79.</p>
[6]	<p><b>"Requirement of pro-sequence for the production of active subtilisin E in <i>Escherichia coli</i>."</b>  Ikemura H., Takagi H., Inouye M.  J. Biol. Chem. 262:7859-7864(1987) [PubMed: 3108260] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-156.  <u>Strain:</u> 168 / PY79.</p>
[7]	<p><b>"Isolation, characterization and structure of subtilisin from a thermostable <i>Bacillus subtilis</i> isolate."</b>  Kamal M., Hoeoeg J.-O., Kaiser R., Shafqat J., Razzaki T., Zaidi Z.H., Joernvall H.  FEBS Lett. 374:363-366(1995) [PubMed: 7589571] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 113-323.  <u>Strain:</u> RT-5.</p>
[8]	<p><b>"<i>Bacillus subtilis</i> subtilisin gene (aprE) is expressed from a sigma A (sigma 43) promoter in vitro and in vivo."</b>  Park S.S., Wong S.L., Wang L.F., Doi R.H.  J. Bacteriol. 171:2657-2665(1989) [PubMed: 2496113] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-13.  <u>Strain:</u> 168.</p>
[9]	<p><b>"Location of the targets of the hpr-97, sacU32(Hy), and sacQ36(Hy) mutations in upstream regions of the subtilisin promoter."</b>  Henner D.J., Ferrari E., Perego M., Hoch J.A.  J. Bacteriol. 170:296-300(1988) [PubMed: 2447063] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-8.</p>
[10]	<p><b>"Determination of the signal peptidase cleavage site in the preprosubtilisin of <i>Bacillus subtilis</i>."</b>  Wong S.L., Doi R.H.  J. Biol. Chem. 261:10176-10181(1986) [PubMed: 3090033] [Abstract]  <u>Cited for:</u> PROTEIN SEQUENCE OF 30-54, SIGNAL PEPTIDE CLEAVAGE SITE.  <u>Strain:</u> 168 / DB104.</p>
[11]	<p><b>"Pro-sequence of subtilisin can guide the refolding of denatured subtilisin in an intermolecular process."</b>  Zhu X.L., Ohta Y., Jordan F., Inouye M.  Nature 339:483-484(1989) [PubMed: 2657436] [Abstract]  <u>Cited for:</u> DOMAIN, MUTAGENESIS OF ASP-138.</p>
[12]	<p><b>"The crystal structure of an autoprocessed Ser221Cys-subtilisin E-propeptide complex at 2.0-A resolution."</b>  Jain S.C., Shinde U., Li Y., Inouye M., Berman H.M.  J. Mol. Biol. 284:137-144(1998) [PubMed: 9811547] [Abstract]</p>

Cited for: X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF MUTANT CYS-327 OF MATURE SUBTILISIN IN COMPLEX WITH PROPEPTIDE AND CALCIUM IONS, COFACTOR.  
Strain: 168.

## Cross-references

### Sequence databases

- ☒ EMBL
- ☐ GenBank
- ☐ DDBJ

K01988 Genomic DNA. Translation: AAA22742.1.  
Y14083 Genomic DNA. Translation: CAA74536.1.  
AL009126 Genomic DNA. Translation: CAB12870.2.  
K01443 Genomic DNA. Translation: AAA22814.1.  
M16639 Genomic DNA. Translation: AAA22744.1.  
M31060 Genomic DNA. Translation: AAA22246.1.  
M19125 Genomic DNA. Translation: AAA22245.1.

PIR

SUBSI. A00972.

RefSeq

NP\_388911.2. NC\_000964.3.

### 3D structure databases

- ☒ PDBe
- ☐ RCSB PDB
- ☐ PDBj

Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
1SCJ	X-ray	2.00	A	107-381	[>]
			B	36-106	[>]

ProteinModelPortal

P04189.

SMR

P04189. Positions 36-381.

DisProt

DP00394.

ModBase

Search...

### Protein family/group databases

MEROPS

I09.001.  
S08.036.

### Genome annotation databases

EnsemblBacteria

EBBACT000000004012; EBBACP000000004012;  
EBBACG000000004004.

GeneID

939313.

GenomeReviews

Gene locus BSU10300 in contig AL009126\_GR.

KEGG

bsu:BSU10300.

NMPDR

fig|224308.1.peg.1030.

### Organism-specific databases

GenoList

BSU10300. [Micado]

CMR

Search...

### Phylogenomic databases

GeneTree

EBGT000500000000194.



HOGENOM	HBG752219.
ProtClustDB	CLSK872792.
<b>Enzyme and pathway databases</b>	
BioCyc	BSUB:BSU10300-MONOMER.
BRENDA	3.4.21.62. 150.
<b>Family and domain databases</b>	
InterPro	IPR000209. Peptidase_S8/S53. IPR022398. Peptidase_S8/S53_AS. IPR015500. Peptidase_S8_subtilisin-rel. IPR009020. Prot_inh_propept. IPR010259. Prot_inh_S8A. [Graphical view]
Gene3D	G3DSA:3.40.50.200. Pept_S8_S53. 1 hit.
PANTHER	PTHR10795. SubtilSerProt. 1 hit.
Pfam	PF05922. Inhibitor_I9. 1 hit. PF00082. Peptidase_S8. 1 hit. [Graphical view]
PRINTS	PR00723. SUBTILISIN.
SUPFAM	SSF52743. Pept_S8_S53. 1 hit. SSF54897. Prot_inh_propept. 1 hit.
PROSITE	PS00136. SUBTILASE_ASP. 1 hit. PS00137. SUBTILASE_HIS. 1 hit. PS00138. SUBTILASE_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Entry information</b>	
Entry name	SUBT_BACSU
Accession	Primary (citable) accession number: <b>P04189</b> Secondary accession number(s): O07613, P70989
Entry history	Integrated into March 20, 1987 UniProtKB/Swiss-Prot: Last sequence update: June 16, 2009 Last modified: February 8, 2011 This is version 108 of the entry and version 3 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Prokaryotic Protein Annotation Program
<b>Relevant documents</b>	
Bacillus subtilis Bacillus subtilis (strain 168): entries, gene names and cross-references to SubtiList	
PDB cross-references Index of Protein Data Bank (PDB) cross-references	

Peptidase families  
Classification of peptidase families and list of entries

SIMILARITY comments  
Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact

EMBL-EBI



## Protein

Translations of Life

Display Settings: GenPept

subtilisin (gtg start codon) [*Bacillus subtilis*]

GenBank: AAA22742.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAA22742 381 aa linear BCT 26-APR-1993  
DEFINITION subtilisin (gtg start codon) [*Bacillus subtilis*].  
ACCESSION AAA22742  
VERSION AAA22742.1 GI:143520  
DBSOURCE locus BACSBTL accession K01988.1  
KEYWORDS  
SOURCE *Bacillus subtilis*  
ORGANISM *Bacillus subtilis*  
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
REFERENCE 1 (residues 1 to 381)  
AUTHORS Stahl, M.L. and Ferrari, E.  
TITLE Replacement of the *Bacillus subtilis* subtilisin structural gene with an in vitro-derived deletion mutation  
JOURNAL J. Bacteriol. 158 (2), 411-418 (1984)  
PUBMED 6427178  
COMMENT [1] reports an in vitro-derived deletion mutant (delta-apr-684), which produces only 10% of wild-type serine protease activity. The deletion is from base 178 to 871.  
Method: conceptual translation.

FEATURES  
source Location/Qualifiers  
1..381  
/organism="Bacillus subtilis"  
/strain="W168"  
/sub\_strain="PY79"  
/db\_xref="taxon:1423"  
Protein 1..381  
/name="subtilisin (gtg start codon)"  
Region 37..104  
/region\_name="Inhibitor\_I9"  
/note="Peptidase inhibitor I9; pfam05922"  
/db\_xref="CDD:147849"  
Region 131..359  
/region\_name="Peptidases\_S8\_Subtilisin\_subset"  
/note="Peptidase S8 family domain in Subtilisin proteins; cd07477"  
/db\_xref="CDD:173803"  
Site order(138,170,213,232,263,327)  
/site\_type="active"  
/db\_xref="CDD:173803"  
Site order(138,170,327)  
/site\_type="other"  
/note="catalytic residues"  
/db\_xref="CDD:173803"  
CDS 1..381  
/coded\_by="K01988.1:137..1282"  
/transl\_table=11

ORIGIN  
1 mrskklwisl lfaltliftm afsnmsaqaa gksstekkyi vgfkqtsam ssakkkdvis  
61 ekggkvqkqf kyvnaaaatl dekaikelkk dpsvayveed hiaheyagsv pygisqikap  
121 alhsqgytgs nvkvavidsg idsshpdlnv rggasfvpse tnpyqdgssh gthvagtiaa  
181 lnnsigvlgv spsaslyavk vldstgsggy swiingiewa isnnmdvinm slggptgsta  
241 lktvvdkavs sgivvaaag negsggstst vgyypakypst iavgavnsen grasfssags  
301 eldvmapgvs iqstlpggty gayngtamat phvagaaali lskhptwtna qvrdrlesta  
361 tylgnsfyyg kglinvqaaa q  
//



## Protein

Translations of Life

Display Settings: GenPept

**subtilisin E protease (GTG start codon; putative); putative [Bacillus subtilis]**

GenBank: AAA22814.1

[FASTA](#) [Graphics](#)

Loading ... 

## Protein

Translations of Life

Display Settings: GenPept

## pre-pro-subtilisin [Bacillus subtilis]

GenBank: AAA22744.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAA22744 156 aa linear BCT 26-APR-1993  
DEFINITION pre-pro-subtilisin [Bacillus subtilis].  
ACCESSION AAA22744  
VERSION AAA22744.1 GI:143523  
DBSOURCE locus BACSBTLA accession M16639.1  
KEYWORDS  
SOURCE Bacillus subtilis  
ORGANISM Bacillus subtilis  
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
REFERENCE 1 (residues 1 to 156)  
AUTHORS Ikemura,H., Takagi,H. and Inouye,M.  
TITLE Requirement of pro-sequence for the production of active subtilisin  
E in Escherichia coli  
JOURNAL J. Biol. Chem. 262 (16), 7859-7864 (1987)  
PUBMED 3108260  
COMMENT Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..156  
/organism="Bacillus subtilis"  
/strain="W168"  
/sub\_strain="PY79"  
/db\_xref="taxon:1423"  
Protein 1..156  
/name="pre-pro-subtilisin"  
sig\_peptide 1..29  
/note="pre-pro-subtilisin signal peptide"  
mat\_peptide 30..>156  
/product="pro-subtilisin"  
Region 37..104  
/region\_name="Inhibitor\_I9"  
/note="Peptidase inhibitor I9; pfam05922"  
/db\_xref="CDD:147849"  
mat\_peptide 107..>156  
/product="subtilisin"  
Region 122..>156  
/region\_name="Peptidases\_S8\_S53"  
/note="Peptidase domain in the S8 and S53 families;  
cl10459"  
/db\_xref="CDD:175216"  
CDS 1..156  
/coded\_by="M16639.1:133..>600"  
/transl\_table=11  
ORIGIN  
1 mrskklwisl lfaltliftm afsnmsaqaa gksstekkyi vgfkqtsam ssakkkdvis  
61 ekggkvqkqf kyvnaaatl dekaavkelkk dpsvayveed hiaheyaqsv pygisqikap  
121 alhsqgytgs nvkvavidsg idsshpdlnv rggasf  
//

**P00781 (SUBD\_BACLI) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified January 11, 2011. Version 76.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Subtilisin DY</b> EC=3.4.21.62
Gene names	Name: <b>apr</b>
Organism	<b>Bacillus licheniformis</b>
Taxonomic identifier	1402 [NCBI]
Taxonomic lineage	Bacteria › Firmicutes › Bacillales › Bacillaceae › Bacillus

**Protein attributes**

Sequence length	274 AA.
Sequence status	Complete.
Protein existence	Evidence at protein level.

**General annotation (Comments)**










Function	Subtilisin is an extracellular alkaline serine protease, it catalyzes the hydrolysis of proteins and peptide amides.
Catalytic activity	Hydrolysis of proteins with broad specificity for peptide bonds, and a preference for a large uncharged residue in P1. Hydrolyzes peptide amides.
Cofactor	Binds 2 calcium ions per subunit.
Subcellular location	Secreted.
Miscellaneous	Secretion of subtilisin is associated with onset of sporulation, and many mutations which block sporulation at early stages affect expression levels of subtilisin. However, subtilisin is not necessary for normal sporulation.
Sequence similarities	Belongs to the peptidase S8 family.

**Ontologies**

<b>Keywords</b>	
Biological process	Sporulation
Cellular component	Secreted
Ligand	Calcium Metal-binding
Molecular function	Hydrolase Protease Serine protease

Technical term	3D-structure Direct protein sequencing
<b>Gene Ontology (GO)</b>	
Biological process	proteolysis Inferred from electronic annotation. Source: InterPro sporulation resulting in formation of a cellular spore Inferred from electronic annotation. Source: UniProtKB-KW
Cellular component	extracellular region Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	metal ion binding Inferred from electronic annotation. Source: UniProtKB-KW serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro
Complete GO annotation...	

**Sequence annotation (Features)**

Feature key	Position (s)	Length	Description	Graphical view	Feature identifier
<b>Molecule processing</b>					
 Chain	1 – 274	274	Subtilisin DY		PRO_0000076417
<b>Sites</b>					
<input type="checkbox"/> Active site	32	1	Charge relay system		
<input type="checkbox"/> Active site	63	1	Charge relay system		
<input type="checkbox"/> Active site	220	1	Charge relay system		
<input type="checkbox"/> Metal binding	2	1	Calcium 1		
<input type="checkbox"/> Metal binding	41	1	Calcium 1		
<input type="checkbox"/> Metal binding	74	1	Calcium 1; via carbonyl oxygen		
<input type="checkbox"/> Metal binding	76	1	Calcium 1		
<input type="checkbox"/> Metal binding	80	1	Calcium 1; via carbonyl oxygen		



<input type="checkbox"/>	Metal binding	168	1	Calcium 2; via carbonyl oxygen		
<input type="checkbox"/>	Metal binding	170	1	Calcium 2; via carbonyl oxygen		
<input type="checkbox"/>	Metal binding	173	1	Calcium 2; via carbonyl oxygen		

Secondary structure

1 ..... 274

■ Helix ■ Strand ■ Turn

Details...

Sequences

Sequence	Length	Mass (Da)
<input checked="" type="checkbox"/> P00781 [UniParc]. Last modified July 21, 1986. Version 1. Checksum: 0154696E22F46533	FASTA 274	27,436
<div><div><div><div>102030405060</div><div>AQTVPYGIPLIKADKVQAQGYKGANVKVGIDTGIAASHTDLKVVGGASFVSGESYNTDG</div></div><div><div>708090100110120</div><div>NGHGTHVAGTVAALDNTTGV LGVAPNVSLY AIKVLNSSGS GTYSAIVSGIEWATQNGLDV</div></div><div><div>130140150160170180</div><div>INMSLGGPSGSTALKQAVDK AYASGIVVVA AAGNSGSSGS QNTIGYPAKY DSVIavgavd</div></div><div><div>190200210220230240</div><div>SNKNRASFS SVGAELVMAP GSVYSTYPS NTYTSLNGTS MASPHVAGAA ALILSKYPTL</div></div><div><div>250260270</div><div>SASQVRNRLS STATNLGDSF YYGKGLINVE AAAQ</div></div></div></div>		

« Hide

References

[1]

"Primary structure of subtilisin DY."  
Nedkov P., Oberthur W., Braunitzer G.  
Hoppe-Seyler's Z. Physiol. Chem. 364:1537-1540(1983) [PubMed: 6420308] [Abstract]  
Cited for: PROTEIN SEQUENCE.  
Strain: DY.

[2]

"Crystal structure of subtilisin DY, a random mutant of subtilisin Carlsberg."  
Eschenburg S., Genov N., Peters K., Fittkau S., Stoeva S., Wilson K.S., Betzel C.  
Eur. J. Biochem. 257:309-318(1998) [PubMed: 9826175] [Abstract]  
Cited for: X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).  
Strain: DY.

**Cross-references****3D structure databases**

<a href="#">PDBe</a> <a href="#">RCSB PDB</a> <a href="#">PDBj</a>	Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
	1BH6	X-ray	1.75	A	1-274	[»]

ProteinModelPortal	P00781.
--------------------	---------

SMR	P00781. Positions 1-274.
-----	--------------------------

ModBase	Search...
---------	-----------

**Protein family/group databases**

MEROPS	S08.037.
--------	----------

**Enzyme and pathway databases**

BRENDA	3.4.21.62. 1017.
--------	------------------

**Family and domain databases**

InterPro	IPR000209. Peptidase_S8/S53. IPR022398. Peptidase_S8/S53_AS. IPR015500. Peptidase_S8_subtilisin-rel. [Graphical view]
----------	--

Gene3D	G3DSA:3.40.50.200. Pept_S8_S53. 1 hit.
--------	--

PANTHER	PTHR10795. SubtilSerProt. 1 hit.
---------	----------------------------------

Pfam	PF00082. Peptidase_S8. 1 hit. [Graphical view]
------	---

PRINTS	PR00723. SUBTILISIN.
--------	----------------------

SUPFAM	SSF52743. Pept_S8_S53. 1 hit.
--------	-------------------------------

PROSITE	PS00136. SUBTILASE_ASP. 1 hit. PS00137. SUBTILASE_HIS. 1 hit. PS00138. SUBTILASE_SER. 1 hit. [Graphical view]
---------	--

ProtoNet	Search...
----------	-----------

**Entry information**

Entry name	SUBD_BACLI
Accession	Primary (citable) accession number: <b>P00781</b>
Entry history	Integrated into July 21, 1986 UniProtKB/Swiss-Prot: Last sequence July 21, 1986 update: Last modified: January 11, 2011 This is version 76 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Prokaryotic Protein Annotation Program

**Relevant documents**

PDB cross-references  
Index of Protein Data Bank (PDB) cross-references

Peptidase families  
Classification of peptidase families and list of entries

SIMILARITY comments  
Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



**P29599 (SUBB\_BACLE) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified January 11, 2011. Version 72.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Subtilisin BL</b> EC=3.4.21.62 <i>Alternative name(s):</i> Alkaline protease
Organism	<b>Bacillus lentus</b>
Taxonomic identifier	1467 [NCBI]
Taxonomic lineage	Bacteria › Firmicutes › Bacillales › Bacillaceae › Bacillus

**Protein attributes**

Sequence length	269 AA.
Sequence status	Complete.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Subtilisin is an extracellular alkaline serine protease, it catalyzes the hydrolysis of proteins and peptide amides.
Catalytic activity	Hydrolysis of proteins with broad specificity for peptide bonds, and a preference for a large uncharged residue in P1. Hydrolyzes peptide amides.
Cofactor	Binds 2 calcium ions per subunit.
Subcellular location	Secreted.
Miscellaneous	Secretion of subtilisin is associated with onset of sporulation, and many mutations which block sporulation at early stages affect expression levels of subtilisin. However, subtilisin is not necessary for normal sporulation.
Sequence similarities	Belongs to the peptidase S8 family.


**Ontologies**

<b>Keywords</b>	
Biological process	Sporulation
Cellular component	Secreted
Ligand	Calcium Metal-binding
Molecular function	Hydrolase Protease Serine protease



Technical term	3D-structure
<b>Gene Ontology (GO)</b>	
Biological process	proteolysis Inferred from electronic annotation. Source: InterPro sporulation resulting in formation of a cellular spore Inferred from electronic annotation. Source: UniProtKB-KW
Cellular component	extracellular region Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	metal ion binding Inferred from electronic annotation. Source: UniProtKB-KW serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro
Complete GO annotation...	

### Sequence annotation (Features)

	Feature key	Position (s)	Length	Description	Graphical view	Feature identifier
<b>Molecule processing</b>						
	Chain	1 – 269	269	Subtilisin BL		PRO_0000076416
<b>Sites</b>						
<input type="checkbox"/>	Active site	32	1	Charge relay system		
<input type="checkbox"/>	Active site	62	1	Charge relay system		
<input type="checkbox"/>	Active site	215	1	Charge relay system		
<input type="checkbox"/>	Metal binding	2	1	Calcium 1		
<input type="checkbox"/>	Metal binding	40	1	Calcium 1		
<input type="checkbox"/>	Metal binding	73	1	Calcium 1; via carbonyl oxygen		
<input type="checkbox"/>	Metal binding	75	1	Calcium 1		
<input type="checkbox"/>	Metal binding	77	1	Calcium 1; via carbonyl oxygen		

<input type="checkbox"/>	Metal binding	79	1	Calcium 1; via carbonyl oxygen		
<input type="checkbox"/>	Metal binding	163	1	Calcium 2; via carbonyl oxygen		
<input type="checkbox"/>	Metal binding	165	1	Calcium 2; via carbonyl oxygen		
<input type="checkbox"/>	Metal binding	168	1	Calcium 2; via carbonyl oxygen		

Secondary structure

1 ..... 269

■ Helix ■ Strand ■ Turn

Details...

Sequences

Sequence	Length	Mass (Da)
<input type="checkbox"/> P29599 [UniParc]. Last modified April 1, 1993. Version 1. Checksum: E8AFF1A6A9E2676B	FASTA 269	26,824
<div><div>102030405060 AQSVPWGISR VQAPAAHNRG LTGSGVKVAV LDTGISTHPD LNIRGGASFV PGEPSTQDGN 708090100110120 GHGTHVAGTI AALNNSIGVL GVAPSAELYA VKVLGADGRG AISSIAQGLE WAGNNGMHVA 130140150160170180 NLSLGSPSPS ATLEQAVNSA TSRGVLVVA A SGNSGASSIS YPARYANAMA VGATDQNNNR 190200210220230240 ASFSQYGAGL DIVAPGVNVQ STYPGSTYAS LNGTSMATPH VAGAAALVKQ KNPSWSNVQI 250260 RNHLKNTATS LGSTNLYGSG LVNAEAATR</div></div>		

« Hide

References

[1] "The crystal structure of the Bacillus lentus alkaline protease, subtilisin BL, at 1.4-A resolution."  
Goddette D.W., Paech C., Yang S.S., Mielenz J.R., Bystroff C., Wilke M.E., Fletterick R.J.  
J. Mol. Biol. 228:580-595(1992) [PubMed: 1453465] [Abstract]  
Cited for: X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS).

## Cross-references

### 3D structure databases

<a href="#">PDBe</a> <a href="#">RCSB PDB</a> <a href="#">PDBj</a>	Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
	1ST3	X-ray	1.40	A	2-269	[»]
ProteinModelPortal	P29599.					
SMR	P29599. Positions 1-269.					
ModBase	Search...					

### Protein family/group databases

MEROPS | S08.003.

### Enzyme and pathway databases

BRENDA | 3.4.21.62. 18569.

### Family and domain databases

InterPro	IPR000209. Peptidase_S8/S53. IPR022398. Peptidase_S8/S53_AS. IPR015500. Peptidase_S8_subtilisin-rel. [Graphical view]
Gene3D	G3DSA:3.40.50.200. Pept_S8_S53. 1 hit.
PANTHER	PTHR10795. SubtilSerProt. 1 hit.
Pfam	PF00082. Peptidase_S8. 1 hit. [Graphical view]
PRINTS	PR00723. SUBTILISIN.
SUPFAM	SSF52743. Pept_S8_S53. 1 hit.
PROSITE	PS00136. SUBTILASE_ASP. 1 hit. PS00137. SUBTILASE_HIS. 1 hit. PS00138. SUBTILASE_SER. 1 hit. [Graphical view]
ProtoNet	Search...

## Entry information

Entry name	SUBB_BACLE
Accession	Primary (citable) accession number: <b>P29599</b>
Entry history	Integrated into April 1, 1993 UniProtKB/Swiss-Prot: Last sequence update: April 1, 1993 Last modified: January 11, 2011 This is version 72 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Prokaryotic Protein Annotation Program

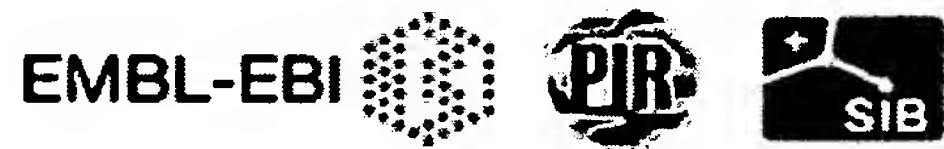
**Relevant documents**

PDB cross-references  
Index of Protein Data Bank (PDB) cross-references

Peptidase families  
Classification of peptidase families and list of entries

SIMILARITY comments  
Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact





**P12544 (GRAA\_HUMAN) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 131.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Granzyme A</b> EC=3.4.21.78 <i>Alternative name(s):</i> CTL tryptase Cytotoxic T-lymphocyte proteinase 1 Fragmentin-1 Granzyme-1 Hanukkah factor Short name=H factor Short name=HF
Gene names	Name: <b>GZMA</b> Synonyms: CTLA3, HFSP
Organism	<b>Homo sapiens (Human)</b> [Complete proteome]
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo

**Protein attributes**

Sequence length	262 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	This enzyme is necessary for target cell lysis in cell-mediated immune responses. It cleaves after Lys or Arg. Cleaves APEX1 after 'Lys-31' and destroys its oxidative repair activity. Involved in apoptosis. <a href="#">Ref.10</a>
Catalytic activity	Hydrolysis of proteins, including fibronectin, type IV collagen and nucleolin. Preferential cleavage: -Arg- -Xaa-, -Lys- -Xaa- >> -Phe- -Xaa- in small molecule substrates.
Subunit structure	Homodimer; disulfide-linked. Interacts with APEX1. <a href="#">Ref.10</a>
Subcellular location	<u>Isoform alpha</u> : Secreted. Cytoplasmic granule.
Induction	Dexamethasone (DEX) induces expression of isoform beta and represses expression of isoform alpha. The alteration in expression is mediated by binding of glucocorticoid receptor to independent promoters adjacent to the alternative first exons of isoform alpha and isoform beta. <a href="#">Ref.5</a>

## Protein

Translations of Life

Display Settings: GenPept

## Hanukah factor serine protease precursor [Homo sapiens]

GenBank: AAA52647.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAA52647 262 aa linear PRI 08-NOV-1994  
DEFINITION Hanukah factor serine protease precursor [Homo sapiens].  
ACCESSION AAA52647  
VERSION AAA52647.1 GI:306845  
DBSOURCE locus HUMHFPSP accession M18737.1  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 262)  
AUTHORS Gershenfeld, H.K., Hershenberger, R.J., Shows, T.B. and Weissman, I.L.  
TITLE Cloning and chromosomal assignment of a human cDNA encoding a T  
cell- and natural killer cell-specific trypsin-like serine protease  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 85 (4), 1184-1188 (1988)  
PUBMED 3257574  
COMMENT On Jul 26, 1993 this sequence version replaced gi:184023.  
Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..262  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="5"  
Protein 1..262  
/name="Hanukah factor serine protease precursor"  
sig\_peptide 1..28  
/note="Hanukah factor serine protease signal peptide"  
mat\_peptide 29..262  
/product="Hanukah factor serine protease"  
Region 29..254  
/region\_name="Tryp\_SpC"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cd00190"  
/db\_xref="CDD:29152"  
Site 29  
/site\_type="cleavage"  
/db\_xref="CDD:29152"  
Site order(69,114,212)  
/site\_type="active"  
/db\_xref="CDD:29152"  
Site order(206,227,229)  
/site\_type="other"  
/note="substrate binding sites"  
/db\_xref="CDD:29152"  
CDS 1..262  
/gene="GJA1P1"  
/coded\_by="M18737.1:1..789"  
/db\_xref="GDB:G00-125-920"  
ORIGIN  
1 mrnsyrflas slsvvsl1l1 ipedvcekii ggnevtphsr pymvllslldr kticagalia  
61 kdwvltaaahc nlnkrsqvil gahsitreep tkqimlvkke fpypcydpat regdlkllql  
121 tekakinkyv tilhlpkkgd dvkpqgtmcqv agwgrthnsa swsdtlrevn itiidrkvcn  
181 drnhynfnpv igmnmvcags lrggrdscng dsqspllceg vfrgvtsfgl enkcqdprrgp  
241 gvyillskkh lnwiimtikg av  
//

## Protein

Translations of Life

Display Settings: GenPept

## granzyme A [Homo sapiens]

GenBank: AAD00009.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAD00009 23 aa linear PRI 25-MAR-1999  
DEFINITION granzyme A [Homo sapiens].  
ACCESSION AAD00009  
VERSION AAD00009.1 GI:4096788  
DBSOURCE locus HSU40006 accession [U40006.1](#)  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 23)  
AUTHORS Goralski, T.J. and Krensky, A.M.  
TITLE The upstream region of the human granzyme A locus contains both  
positive and negative transcriptional regulatory elements  
J. Immunol. (1995) In press  
REFERENCE 2 (residues 1 to 23)  
AUTHORS Goralski, T.J. and Krensky, A.M.  
TITLE Direct Submission  
JOURNAL Submitted (03-NOV-1995) T.J. Goralski, Pediatrics, Stanford  
University School of Medicine, Stanford, CA 94305, USA  
COMMENT Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..23  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
Protein 1..23  
/product="granzyme A"  
CDS 1..23  
/coded\_by="U40006.1:1541..>1610"  
/note="serine protease"  
ORIGIN  
1 mrnsyrflas slsvvvslll ipe  
//

**P11032 (GRAA\_MOUSE) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified March 8, 2011. Version 119.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Granzyme A</b> EC=3.4.21.78 <i>Alternative name(s):</i> Autocrine thymic lymphoma granzyme-like serine protease CTLA-3 Fragmentin-1 T cell-specific serine protease 1 Short name=TSP-1
Gene names	Name: <b>Gzma</b> Synonyms: Ctla-3, Ctla3, Mtsp-1
Organism	<b>Mus musculus (Mouse)</b>
Taxonomic identifier	10090 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires › Rodentia › Sciurognathi › Muroidea › Muridae › Murinae › Mus › Mus

**Protein attributes**

Sequence length	260 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	This enzyme is necessary for target cell lysis in cell-mediated immune responses. It cleaves after Lys or Arg. Cleaves APEX1 after 'Lys-30' and destroys its oxidative repair activity. Involved in apoptosis <span>By similarity</span> .
Catalytic activity	Hydrolysis of proteins, including fibronectin, type IV collagen and nucleolin. Preferential cleavage: -Arg- -Xaa-, -Lys- -Xaa->> -Phe- -Xaa- in small molecule substrates.
Subunit structure	Interacts with APEX1 <span>By similarity</span> . Homodimer; disulfide-linked.
Subcellular location	Secreted. Cytoplasmic granule. Note: Cytoplasmic granules of cytolytic T-lymphocytes.
Tissue specificity	Found in cytotoxic lymphocytes and in normal lymphoid tissues such as thymus and spleen.
Sequence similarities	Belongs to the peptidase S1 family. Granzyme subfamily. Contains 1 peptidase S1 domain.



**Caution**

The predicted cleavage site for the activation peptide of HF2 is uncertain. It could have either 2 (ER) or 7 (KRGGCER) AA.

**Ontologies****Keywords**

Biological process	Apoptosis Cytolysis
Cellular component	Secreted
Coding sequence diversity	Alternative splicing
Domain	Signal
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Glycoprotein Zymogen
Technical term	Direct protein sequencing

**Gene Ontology (GO)**

Biological process	<p>cleavage of lamin Inferred from sequence or structural similarity. Source: UniProtKB</p> <p>cytolysis Inferred from electronic annotation. Source: UniProtKB-KW</p> <p>negative regulation of DNA binding Inferred from sequence or structural similarity. Source: UniProtKB</p> <p>negative regulation of endodeoxyribonuclease activity Inferred from sequence or structural similarity. Source: UniProtKB</p> <p>negative regulation of oxidoreductase activity Inferred from sequence or structural similarity. Source: UniProtKB</p> <p>positive regulation of apoptosis Inferred from sequence or structural similarity. Source: UniProtKB</p>
Cellular component	<p>extracellular region Inferred from electronic annotation. Source: UniProtKB-SubCell</p>
Molecular function	<p>protein homodimerization activity Inferred from sequence or structural similarity. Source: UniProtKB</p> <p>serine-type endopeptidase activity Inferred from sequence or structural similarity. Source: UniProtKB</p>

Complete GO annotation...

## Alternative products

This entry describes 2 isoforms produced by **alternative splicing**. [Align] [Select]

### Isoform HF1 (identifier: **P11032-1**)

*This isoform has been chosen as the 'canonical' sequence. All positional information in this entry refers to it. This is also the sequence that appears in the downloadable versions of the entry.*














Note: More abundant in lymphoid tissues than isoform HF2.

### Isoform HF2 (identifier: **P11032-2**)

*The sequence of this isoform differs from the canonical sequence as follows:*

1-23: MRNASGPRGPSLATLLFLLIPE → MSKEMNEILLSWEINLSSKR

## Sequence annotation (Features)

	Feature key	Position(s)	Length	Description	Graphical view
<b>Molecule processing</b>					
	Signal peptide	1 – 26	26		
	Propeptide	27 – 28	2	Activation peptide	
	Chain	29 – 260	232	Granzyme A	
<b>Regions</b>					
	Domain	29 – 257	229	Peptidase S1	
<b>Sites</b>					
	Active site	69	1	Charge relay system By similarity	
	Active site	113	1	Charge relay system By similarity	
	Active site	211	1	Charge relay system By similarity	
<b>Amino acid modifications</b>					
	Glycosylation	157	1	N-linked (GlcNAc...) Potential	
	Glycosylation	169	1	N-linked (GlcNAc...) Potential	
	Disulfide bond	54 ↔ 70		By similarity	
	Disulfide bond	147 ↔ 217		By similarity	
	Disulfide bond	178 ↔ 196		By similarity	
	Disulfide bond	207 ↔ 232		By similarity	

## Natural variations

<input checked="" type="checkbox"/>	Alternative sequence	1 – 23	23	MRNAS...LLIPE → MSKEMNEILLSWEINLSSKR in isoform HF2.	
-------------------------------------	----------------------	--------	----	--	--

## Sequences

Sequence		Length	Mass (Da)
<input checked="" type="checkbox"/> Isoform HF1 [UniParc].	FASTA	260	28,599

Last modified July 1, 1993. Version 2.  
Checksum: 7705352F08DEED8D

```

      10      20      30      40      50      60
MRNASGPRGP SLATLLFLLL IPEGGCERII GGDTVVPHSR PYMALLKLSS NTICAGALIE

      70      80      90     100     110     120
KNWVLTAHC NVGKRSKFIL GAHSINKEPE QQILTVKKAF PYPCYDEYTR EGDQLVRLK

     130     140     150     160     170     180
KKATVNRNVÄ ILHLPKKGDD VKPGTRCRVÄ GWGRFGNKSÄ PSETLREVNI TVIDRKICND

     190     200     210     220     230     240
EKHYNFHPVI GLNMICAGDL RGGKDSCNGD SGSPLLCDGI LRGITSFGGE KCGDRRWPGV

     250     260
YTFLSDKHLN WIKKIMKGSV

```

« Hide

<input checked="" type="checkbox"/> Isoform HF2.	FASTA	257	28,541
Checksum: 7D67F97E532F7440			
Show »			

## References




« Hide 'large scale' references

- |     |  |
|-----|--|
| [1] | <p><b>"cDNA clones from autocrine thymic lymphoma cells encode two mitogenic proteins, a serine protease and a truncated T-cell receptor beta-chain."</b><br/>           Bogenberger J., Haas M.<br/>           Oncogene Res. 3:301-312(1988) [PubMed: 2976140] [Abstract]<br/> <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].<br/> <u>Strain:</u> C57BL/6.</p>                           |
| [2] | <p><b>"Organization of the gene encoding the mouse T-cell-specific serine proteinase 'granzyme A'."</b><br/>           Ebnet K., Kramer M.D., Simon M.M.<br/>           Genomics 13:502-508(1992) [PubMed: 1639378] [Abstract]<br/> <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA].<br/> <u>Strain:</u> C57BL/6.</p>  |
| [3] | <p><b>"Genomic organization of the mouse granzyme A gene. Two mRNAs encode the same mature granzyme A with different leader peptides."</b><br/>           Hershberger R.J., Gershenfeld H.K., Weissman I.L., Su L.<br/>           J. Biol. Chem. 267:25488-25493(1992) [PubMed: 1460043] [Abstract]<br/> <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA] (ISOFORMS HF1 AND HF2).</p> |

[4]	<p><b>"The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC)."</b>  The MGC Project Team  Genome Res. 14:2121-2127(2004) [PubMed: 15489334] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM HF1).</p>
[5]	<p><b>"Granzymes, a family of serine proteases released from granules of cytolytic T lymphocytes upon T cell receptor stimulation."</b>  Jenne D.E., Tschopp J.  Immunol. Rev. 103:53-71(1988) [PubMed: 3292396] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] OF 1-37.</p>
[6]	<p><b>"Cloning of a cDNA for a T cell-specific serine protease from a cytotoxic T lymphocyte."</b>  Gershenfeld H.K., Weissman I.L.  Science 232:854-858(1986) [PubMed: 2422755] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] OF 12-260.</p>
[7]	<p><b>"A family of serine esterases in lytic granules of cytolytic T lymphocytes."</b>  Masson D., Tschopp J.  Cell 49:679-685(1987) [PubMed: 3555842] [Abstract]  <u>Cited for:</u> PROTEIN SEQUENCE OF 29-48.</p>
[8]	<p><b>"Identification of granzyme A isolated from cytotoxic T-lymphocyte-granules as one of the proteases encoded by CTL-specific genes."</b>  Masson D., Zamai M., Tschopp J.  FEBS Lett. 208:84-88(1986) [PubMed: 3533635] [Abstract]  <u>Cited for:</u> PROTEIN SEQUENCE OF 29-53.</p>
[9]	<p><b>"Induction of T cell serine proteinase 1 (TSP-1)-specific mRNA in mouse T lymphocytes."</b>  Simon H.G., Fruth U., Eckerskorn C., Lottspeich F., Kramer M.D., Nerz G., Simon M.M.  Eur. J. Immunol. 18:855-861(1988) [PubMed: 3260181] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] OF 29-46.</p>
+	Additional computationally mapped references.

## Cross-references

### Sequence databases

<ul style="list-style-type: none"> <li>⊗ EMBL</li> <li>⊗ GenBank</li> <li>⊗ DDBJ</li> </ul>	X14799 mRNA. Translation: CAA32905.1. X62542  X60311 Genomic DNA. Translation: CAA44426.1. L01429  L01441 Genomic DNA. Translation: AAA99898.1. L01429  L01441 Genomic DNA. Translation: AAA99897.1. BC061146 mRNA. Translation: AAH61146.1. M26183 mRNA. Translation: AAA37735.1. M13226 mRNA. Translation: AAA40134.1.
---	---

IPI	IPI00111385. IPI00229002.
PIR	A24807. A45061. B45061.
RefSeq	NP_034500.1. NM_010370.2.
UniGene	Mm.15510.

### 3D structure databases

ProteinModelPortal	P11032.
--------------------	---------



SMR	P11032. Positions 29-258.
ModBase	Search...
<b>Protein-protein interaction databases</b>	
STRING	P11032.
<b>Protein family/group databases</b>	
MEROPS	S01.135.
<b>Proteomic databases</b>	
PRIDE	P11032.
<b>Genome annotation databases</b>	
Ensembl	ENSMUST00000023897; ENSMUSP00000023897; ENSMUSG00000023132. ENSMUST00000109243; ENSMUSP00000104866; ENSMUSG00000023132.
GeneID	14938.
KEGG	mmu:14938.
UCSC	uc007rxc.1. mouse.
<b>Organism-specific databases</b>	
CTD	14938.
MGI	MGI:109266. Gzma.
<b>Phylogenomic databases</b>	
eggNOG	roNOG11241.
GeneTree	ENSGT00580000081367.
HOGENOM	HBG755338.
HOVERGEN	HBG013304.
InParanoid	P11032.
OMA	KEHDIML.
OrthoDB	EOG4SXNDS.
PhylomeDB	P11032.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.21.78. 244.
<b>Gene expression databases</b>	
ArrayExpress	P11032.
Bgee	P11032.
CleanEx	MM_GZMA.
Genevestigator	P11032.

GermOnline	ENSMUSG00000023132. Mus musculus.
<b>Family and domain databases</b>	
InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
NextBio	287259.
SOURCE	Search...
<b>Entry information</b>	
Entry name	GRAA_MOUSE
Accession	Primary (citable) accession number: <b>P11032</b> Secondary accession number(s): P15118
Entry history	Integrated into July 1, 1989 UniProtKB/Swiss-Prot: Last sequence July 1, 1993 update: Last modified: March 8, 2011 This is version 119 of the entry and version 2 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Relevant documents</b>	
MGD cross-references Mouse Genome Database (MGD) cross-references in UniProtKB/Swiss-Prot	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact

EMBL-EBI



**P10144 (GRAB\_HUMAN)** ★ Reviewed, UniProtKB/Swiss-Prot

Last modified April 5, 2011. Version 132.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Granzyme B</b> EC=3.4.21.79 <i>Alternative name(s):</i> C11 CTLA-1 Cathepsin G-like 1 Short name=CTSG1 Cytotoxic T-lymphocyte proteinase 2 Short name=Lymphocyte protease Fragmentin-2 Granzyme-2 Human lymphocyte protein Short name=HLP SECT T-cell serine protease 1-3E
Gene names	Name: <b>GZMB</b> Synonyms: CGL1, CSPB, CTLA1, GRB
Organism	<b>Homo sapiens (Human)</b> [Complete proteome]
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo

**Protein attributes**

Sequence length	247 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	This enzyme is necessary for target cell lysis in cell-mediated immune responses. It cleaves after Asp. Seems to be linked to an activation cascade of caspases (aspartate-specific cysteine proteases) responsible for apoptosis execution. Cleaves caspase-3, -7, -9 and 10 to give rise to active enzymes mediating apoptosis.
Catalytic activity	Preferential cleavage: -Asp- -Xaa- >> -Asn- -Xaa- > -Met- -Xaa- , -Ser- -Xaa-. [Ref.12]
Enzyme regulation	Inactivated by the serine protease inhibitor diisopropylfluorophosphate. [Ref.12]



Subcellular location	Cytoplasmic granule. Note: Cytoplasmic granules of cytolytic T-lymphocytes and natural killer cells. <a href="#">Ref.12</a>
Induction	By staphylococcal enterotoxin A (SEA) in peripheral blood leukocytes. <a href="#">Ref.12</a>
Sequence similarities	Belongs to the peptidase S1 family. Granzyme subfamily. Contains 1 peptidase S1 domain.
<b>Ontologies</b>	
<b>Keywords</b>	
Biological process	Apoptosis Cytolysis
Coding sequence diversity	Polymorphism
Domain	Signal
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Glycoprotein Zymogen
Technical term	3D-structure Complete proteome Direct protein sequencing
<b>Gene Ontology (GO)</b>	
Biological process	activation of pro-apoptotic gene products Traceable author statement. Source: Reactome  cleavage of lamin Inferred from direct assay. Source: UniProtKB  cytolysis Inferred from electronic annotation. Source: UniProtKB-KW  induction of apoptosis by intracellular signals Traceable author statement. Source: Reactome
Cellular component	cytosol Inferred from experiment. Source: Reactome  immunological synapse Traceable author statement. Source: UniProtKB  nucleus Traceable author statement. Source: UniProtKB
Molecular function	protein binding Inferred from physical interaction. Source: UniProtKB  serine-type endopeptidase activity Traceable author statement. Source: UniProtKB
Complete GO annotation...	

Binary interactions

With	Entry	#Exp.	IntAct	Notes
PRF1	P14222	2	EBI-2505785,EBI-724466	
SRGN	P10124	1	EBI-2505785,EBI-744915	

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature
-------------	-------------	--------	-------------	----------------	---------

Molecule processing

<input checked="" type="checkbox"/>	Signal peptide	1 – 18	18	Ref.14		
<input type="checkbox"/>	Propeptide	19 – 20	2	Activation peptide		PRO_0000000000
<input checked="" type="checkbox"/>	Chain	21 – 247	227	Granzyme B		PRO_0000000000

Regions

<input checked="" type="checkbox"/>	Domain	21 – 245	225	Peptidase S1		
-------------------------------------	--------	----------	-----	--------------	--	--

Sites

<input type="checkbox"/>	Active site	64	1	Charge relay system		
<input type="checkbox"/>	Active site	108	1	Charge relay system		
<input type="checkbox"/>	Active site	203	1	Charge relay system		

Amino acid modifications

<input type="checkbox"/>	Glycosylation	71	1	N-linked (GlcNAc...)		
<input type="checkbox"/>	Glycosylation	104	1	N-linked (GlcNAc...) Potential		
<input type="checkbox"/>	Disulfide bond	49 ↔ 65				
<input type="checkbox"/>	Disulfide bond	142 ↔ 209				
<input type="checkbox"/>	Disulfide bond	173 ↔ 188				

Natural variations

<input type="checkbox"/>	Natural variant	55	1	R → Q. [dbSNP:rs8192917] Ref.1 Ref.3 Ref.4 Ref.5 Ref.6 Ref.8 Ref.17		VAR_0000000000
<input type="checkbox"/>	Natural variant	94	1	P → A. [dbSNP:rs11539752] Ref.6 Ref.2		VAR_0000000000

<input type="checkbox"/>	Natural variant	247	1	Y → H. [dbSNP:rs2236338] Ref.8 Ref.17		VAR_0
--------------------------	-----------------	-----	---	---	--	-------

Experimental info

<input type="checkbox"/>	Sequence conflict	32 – 33	2	RP → PR AA sequence Ref.12		
<input type="checkbox"/>	Sequence conflict	72	1	V → G in AAA52118. Ref.3		
<input type="checkbox"/>	Sequence conflict	212	1	V → C in AAB59528. Ref.4		

Secondary structure

1 . . . . .

■ Helix

■ Strand

■ Turn

Details...

Sequences

Sequence

Length

Mass (Da)

☐

P10144 [UniParc].  
Last modified January 11, 2011. Version 2.  
Checksum: C652271918EF24F9

FASTA24727,716

102030405060

MQPILLLLAF LLLPRADAGE IIGGHEAKPH SRPYMAYLMI WDQSLKRCG GFLIRDDFVL

708090100110120

TAAHCWGSSI NVTLGAHNIK EQEPTQQFIP VKRPIHPAY NPKNFSNDIM LLQLERKAKR

130140150160170180

TRAVQPLRLP SNKAQVKPGQ TCSVAGWGQT APLGKHSHTL QEVKMTVQED RKCESDLRHY

190200210220230240

YDSTIELCVG DPEIKKTSFK GDSGGPLVCN KVAQGIVSYG RNNGMPPRAC TKVSSFVHWI

KKTMKRY

« Hide

References

« Hide 'large scale' references

[1]

"Induction of mRNA for a serine protease and a beta-thromboglobulin-like protein in mitogen-stimulated human leukocytes."  
Schmid J., Weissmann C.  
J. Immunol. 139:250-256(1987) [PubMed: 2953813] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA], VARIANT GLN-55.

[2]

"Structure and differential mechanisms of regulation of expression of a serine esterase gene in activated human T lymphocytes."

	<p>Caputo A., Fahey D., Lloyd C., Vozab R., McCairns E., Rowe P.B.  J. Biol. Chem. 263:6363-6369(1988) [PubMed: 3258865] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA], VARIANT ALA-94.</p>
[3]	<p><b>"Molecular cloning of an inducible serine esterase gene from human cytotoxic lymphocytes."</b>  Trapani J.A., Klein J.L., White P.C., Dupont B.  Proc. Natl. Acad. Sci. U.S.A. 85:6924-6928(1988) [PubMed: 3261871] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA], VARIANT GLN-55.</p>
[4]	<p><b>"Genomic organization and chromosomal assignment for a serine protease gene (CSPB) expressed by human cytotoxic lymphocytes."</b>  Klein J.L., Shows T.B., Dupont B., Trapani J.A.  Genomics 5:110-117(1989) [PubMed: 2788607] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA], VARIANT GLN-55.</p>
[5]	<p><b>"Nucleotide sequence and genomic organization of a human T lymphocyte serine protease gene."</b>  Caputo A., Sauer D.E., Rowe P.B.  J. Immunol. 145:737-744(1990) [PubMed: 2365998] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA], VARIANT GLN-55.</p>
[6]	<p><b>"Structural organization of the hCTLA-1 gene encoding human granzyme B."</b>  Haddad P., Clement M.-V., Bernard O., Larsen C.-J., Degos L., Sasportes M., Mathieu-Mahul D.  Gene 87:265-271(1990) [PubMed: 2332171] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA], VARIANTS GLN-55 AND ALA-94.</p>
[7]	<p><b>"The DNA sequence and analysis of human chromosome 14."</b>  Heilig R., Eckenberg R., Petit J.-L., Fonknechten N., Da Silva C., Cattolico L., Levy M., Barbe V., De Berardinis V., Ureta-Vidal A., Pelletier E., Vico V., Anthouard V., Rowen L., Madan A., Qin S., Sun H., Du H., Weissenbach J.  Nature 421:601-607(2003) [PubMed: 12508121] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].</p>
[8]	<p><b>"The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC)."</b>  The MGC Project Team  Genome Res. 14:2121-2127(2004) [PubMed: 15489334] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], VARIANTS GLN-55 AND HIS-247.  <u>Tissue:</u> Pancreas.</p>
[9]	<p><b>"Isolation of a cDNA clone encoding a novel form of granzyme B from human NK cells and mapping to chromosome 14."</b>  Dahl C.A., Bach F.H., Chan W., Huebner K., Russo G., Croce C.M., Herfurth T., Cairns J.S.  Hum. Genet. 84:465-470(1990) [PubMed: 2323780] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] OF 1-23.</p>
[10]	<p><b>"Characterization of three serine esterases isolated from human IL-2 activated killer cells."</b>  Hameed A., Lowrey D.M., Lichtenheld M., Podack E.R.  J. Immunol. 141:3142-3147(1988) [PubMed: 3262682] [Abstract]  <u>Cited for:</u> PROTEIN SEQUENCE OF 21-40, CHARACTERIZATION.</p>
[11]	<p><b>"Characterization of granzymes A and B isolated from granules of cloned human cytotoxic T lymphocytes."</b>  Kraehenbuhl O., Rey C., Jenne D.E., Lanzavecchia A., Groscurth P., Carrel S., Tschopp J.  J. Immunol. 141:3471-3477(1988) [PubMed: 3263427] [Abstract]  <u>Cited for:</u> PROTEIN SEQUENCE OF 21-40, CHARACTERIZATION.</p>
[12]	<p><b>"Human granzyme B degrades aggrecan proteoglycan in matrix synthesized by chondrocytes."</b>  Froelich C.J., Zhang X., Turbov J., Hudig D., Winkler U., Hanna W.L.</p>



	J. Immunol. 151:7161-7171(1993) [PubMed: 8258716] [Abstract] <u>Cited for:</u> PROTEIN SEQUENCE OF 21-39, CATALYTIC ACTIVITY, ENZYME REGULATION, SUBCELLULAR LOCATION. <u>Tissue:</u> Lymphocyte.
[13]	<b>"Human cytotoxic lymphocyte granzyme B. Its purification from granules and the characterization of substrate and inhibitor specificity."</b> Poe M., Blake J.T., Boulton D.A., Gammon M., Sigal N.H., Wu J.K., Zweerink H.J. J. Biol. Chem. 266:98-103(1991) [PubMed: 1985927] [Abstract] <u>Cited for:</u> PROTEIN SEQUENCE OF 21-38.
[14]	<b>"Signal peptide prediction based on analysis of experimentally verified cleavage sites."</b> Zhang Z., Henzel W.J. Protein Sci. 13:2819-2824(2004) [PubMed: 15340161] [Abstract] <u>Cited for:</u> PROTEIN SEQUENCE OF 19-33.
[15]	<b>"Crystal structure of the caspase activator human granzyme B, a proteinase highly specific for an Asp-P1 residue."</b> Estebanez-Perpina E., Fuentes-Prior P., Belorgey D., Braun M., Kiefersauer R., Maskos K., Huber R., Rubin H., Bode W. Biol. Chem. 381:1203-1214(2000) [PubMed: 11209755] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF 21-247.
[16]	<b>"The three-dimensional structure of human granzyme B compared to caspase-3, key mediators of cell death with cleavage specificity for aspartic acid in P1."</b> Rotonda J., Garcia-Calvo M., Bull H.G., Geissler W.M., McKeever B.M., Willoughby C.A., Thornberry N.A., Becker J.W. Chem. Biol. 8:357-368(2001) [PubMed: 11325591] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 21-247.
[17]	<b>"Catalog of 680 variations among eight cytochrome p450 (CYP) genes, nine esterase genes, and two other genes in the Japanese population."</b> Saito S., Iida A., Sekine A., Kawauchi S., Higuchi S., Ogawa C., Nakamura Y. J. Hum. Genet. 48:249-270(2003) [PubMed: 12721789] [Abstract] <u>Cited for:</u> VARIANTS GLN-55 AND HIS-247.
+	Additional computationally mapped references.

## Cross-references

### Sequence databases

<ul style="list-style-type: none"> <li>⊗ EMBL</li> <li>⊗ GenBank</li> <li>⊗ DDBJ</li> </ul>	M17016 mRNA. Translation: AAA36627.1. J03189 mRNA. Translation: AAA36603.1. J04071 mRNA. Translation: AAA52118.1. J03072 Genomic DNA. Translation: AAB59528.1. M38193 Genomic DNA. Translation: AAA67124.1. M28879 Genomic DNA. Translation: AAA75490.1. AL136018 Genomic DNA. No translation available. BC030195 mRNA. Translation: AAH30195.1.
IPI	IPI00744785.
PIR	A61021.
RefSeq	NP_004122.2. NM_004131.4.
UniGene	Hs.1051.

**3D structure databases**

<input type="radio"/> PDBe <input type="radio"/> RCSB PDB <input type="radio"/> PDBj	Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
	1FQ3	X-ray	3.10	A/B	21-247	[»]
	1IAU	X-ray	2.00	A	21-247	[»]

ProteinModelPortal

P10144.

ModBase

Search...

**Protein-protein interaction databases**

IntAct

P10144. 5 interactions.

MINT

MINT-4528791.

STRING

P10144.

**Protein family/group databases**

MEROPS

S01.010.

**Proteomic databases**

PRIDE

P10144.

**Genome annotation databases**

Ensembl

ENST00000216341; ENSP00000216341; ENSG00000100453.  
ENST00000382542; ENSP00000371982; ENSG00000100453.  
ENST00000415355; ENSP00000387385; ENSG00000100453.

GeneID

3002.

UCSC

uc001wps.1. human.

**Organism-specific databases**

GeneCards

GC14M005214.

H-InvDB

HIX0011578.

HGNC

HGNC:4709. GZMB.

HPA

CAB000376.  
HPA003418.

MIM

123910. gene.

neXtProt

NX\_P10144.

GenAtlas

Search...

**Phylogenomic databases**

eggNOG

prNOG07952.

HOVERGEN

HBG013304.

InParanoid

P10144.

OrthoDB

EOG4KWJTW.

**Enzyme and pathway databases**

BRENDA	3.4.21.79. 247.
Pathway_Interaction_DB	caspase_pathway. Caspase cascade in apoptosis. cd8tcrdownstreampathway. Downstream signaling in naive CD8+ T cells. il12_2pathway. IL12-mediated signaling events.
Reactome	REACT_578. Apoptosis.

**Gene expression databases**

ArrayExpress	P10144.
Bgee	P10144.
CleanEx	HS_GZMB.
Genevestigator	P10144.
GermOnline	ENSG00000100453. Homo sapiens.

**Family and domain databases**

InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...

**Other Resources**

NextBio	11904.
PMAP-CutDB	P10144.
SOURCE	Search...

**Entry information**

Entry name	GRAB_HUMAN
Accession	Primary (citable) accession number: <b>P10144</b> Secondary accession number(s): Q8N1D2, Q9UCC1
Entry history	Integrated into July 1, 1989 UniProtKB/Swiss-Prot: Last sequence update: January 11, 2011

	Last modified: April 5, 2011 This is version 132 of the entry and version 2 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
Disclaimer	Any medical or genetic information present in this entry is provided for research, educational and informational purposes only. It is not in any way intended to be used as a substitute for professional medical advice, diagnosis, treatment or care.
<b>Relevant documents</b>	
Human chromosome 14 Human chromosome 14: entries, gene names and cross-references to MIM	
Human entries with polymorphisms or disease mutations List of human entries with polymorphisms or disease mutations	
Human polymorphisms and disease mutations Index of human polymorphisms and disease mutations	
MIM cross-references Online Mendelian Inheritance in Man (MIM) cross-references in UniProtKB/Swiss-Prot	
PDB cross-references Index of Protein Data Bank (PDB) cross-references	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact





## Protein

Translations of Life

Display Settings: GenPept

## serine esterase [Homo sapiens]

GenBank: AAA52118.1

[FASTA](#) [Graphics](#)Go to:

LOCUS AAA52118 247 aa linear PRI 01-NOV-1994  
DEFINITION serine esterase [Homo sapiens].  
ACCESSION AAA52118  
VERSION AAA52118.1 GI:181137  
DBSOURCE locus HUMCSE accession J04071.1  
KEYWORDS .  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 247)  
AUTHORS Trapani, J.A., Klein, J.L., White, P.C. and Dupont, B.  
TITLE Molecular cloning of an inducible serine esterase gene from human  
cytotoxic lymphocytes  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 85 (18), 6924-6928 (1988)  
PUBMED 3261871  
COMMENT Draft entry and clean copy of sequence for [1] kindly provided by  
J.A.Trapani, 08-SEP-1988.  
Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..247  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="14q11.2"  
Protein 1..247  
/product="serine esterase"  
sig peptide 1..20  
/note="MHC serine esterase signal peptide"  
mat peptide 21..247  
/product="MHC serine esterase"  
Region 21..243  
/region\_name="Tryp\_Spc"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cd00190"  
/db\_xref="CDD:29152"  
Site 21  
/site\_type="cleavage"  
/db\_xref="CDD:29152"  
Site order(64,108,203)  
/site\_type="active"  
/db\_xref="CDD:29152"  
Site order(197,218,220)  
/site\_type="other"  
/note="substrate binding sites"  
/db\_xref="CDD:29152"  
CDS 1..247  
/gene="CTLA1"  
/coded\_by="J04071.1:17..760"  
/db\_xref="GDB:G00-120-744"  
ORIGIN  
1 mqpilllllaf lllpradage iiggheakph srpymaylmi wdqkslkrsg gfliqddfv1  
61 taahcwgsi ngtlgahnk egeptqqfip vkrpiphpay npknfsndim llqlerkakr  
121 travqplrlp snkaqvkpgg tcsvagwgqt aplgkhshtl qevkmtvqed rkcesdlrhy  
181 ydstielcvg dpeikktsfk gdsggplvcn kvaqgvisyg rnngmpprac tkvssfvhwi  
241 kktmkry  
//

## Protein

Translations of Life

Display Settings: GenPept

## serine protease B [Homo sapiens]

GenBank: AAB59528.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAB59528 247 aa linear PRI 08-AUG-1995  
DEFINITION serine protease B [Homo sapiens].  
ACCESSION AAB59528  
VERSION AAB59528.1 GI:181159  
DBSOURCE locus HUMCSPB accession J03072.1  
KEYWORDS .  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 247)  
AUTHORS Klein, J.L., Shows, T.B., Dupont, B. and Trapani, J.A.  
TITLE Genomic organization and chromosomal assignment for a serine  
protease gene (CSPB) expressed by human cytotoxic lymphocytes  
JOURNAL Genomics 5 (1), 110-117 (1989)  
PUBMED 2788607  
COMMENT Draft entry and computer-readable sequence for [1] kindly provided  
by J.Klein, 29-MAR-1989.  
Aberrant splicing of the mRNA at position 1143 often takes place  
and produces a nonfunctional protein.  
Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..247  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="14q11.2"  
/clone="lambda-GP26."  
Protein 1..247  
/name="serine protease B"  
Region 21..243  
/region\_name="Tryp\_SpC"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cd00190"  
/db\_xref="CDD:29152"  
Site 21  
/site\_type="cleavage"  
/db\_xref="CDD:29152"  
Site order(64,108,203)  
/site\_type="active"  
/db\_xref="CDD:29152"  
Site order(197,218,220)  
/site\_type="other"  
/note="substrate binding sites"  
/db\_xref="CDD:29152"  
CDS 1..247  
/gene="CTLA1"  
/coded\_by="join(J03072.1:233..287,J03072.1:1290..1437,  
J03072.1:1904..2039,J03072.1:2247..2507,  
J03072.1:3148..3291)"  
/db\_xref="GDB:G00-120-744"

ORIGIN  
1 mqpilllllaf lllpradage iiggheakph srpymaylmi wdqkslkrcg gfliqddfv1  
61 taahcwgsi nvtlgahnik egeptqqfip vkrpiphpay npknfsndim llqlerkakr  
121 travqplrlp snkaqvkgqg tcsvagwgqt aplgkhshtl qevkmtvqed rkcesdlrhy  
181 ydstielcvg dpeiktsfk gdsggplvcn kcaqgvisyg rnngmpprac tkvssfvhwi  
241 kktmkry

//

## Protein

Translations of Life

Display Settings: GenPept

## serine protease-like protein precursor [Homo sapiens]

GenBank: AAA36627.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAA36627 247 aa linear PRI 03-AUG-1993  
DEFINITION serine protease-like protein precursor [Homo sapiens].  
ACCESSION AAA36627  
VERSION AAA36627.1 GI:338296  
DBSOURCE locus HUMSP13E accession M17016.1  
KEYWORDS .  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 247)  
AUTHORS Schmid, J. and Weissmann, C.  
TITLE Induction of mRNA for a serine protease and a  
beta-thromboglobulin-like protein in mitogen-stimulated human  
leukocytes  
JOURNAL J. Immunol. 139 (1), 250-256 (1987)  
PUBMED 2953813  
COMMENT Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..247  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
Protein 1..247  
/name="serine protease-like protein precursor"  
sig\_peptide 1..20  
/note="serine protease-like protein signal peptide"  
mat\_peptide 21..247  
/product="serine protease-like protein"  
Region 21..243  
/region\_name="Tryp\_SpC"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cd00190"  
/db\_xref="CDD:29152"  
Site 21  
/site\_type="cleavage"  
/db\_xref="CDD:29152"  
Site order(64,108,203)  
/site\_type="active"  
/db\_xref="CDD:29152"  
Site order(197,218,220)  
/site\_type="other"  
/note="substrate binding sites"  
/db\_xref="CDD:29152"  
CDS 1..247  
/coded\_by="M17016.1:8..751"  
ORIGIN  
1 mqpilllllaf lllpradage iiggheakph srpymaylmi wdqkslkrcg gfliqddfv1  
61 taahcwgsi nvtlgahnk egeptqqfip vkrpiphpay npknfsndim llqlerkakr  
121 travqplrlp snkaqvkgpg tcsvagwggt aplgkhshtl qevkmtvqed rkcesdlrhy  
181 ydstielcvg dpeikktsfk gdsaggplvcn kvaqgivsyt rnngmpprac tkvssfvhwi  
241 kktmkry  
//

## Protein

Translations of Life

Display Settings: GenPept

## cytotoxic T-lymphocyte-associated serine esterase 1 [Homo sapiens]

GenBank: AAA67124.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAA67124 247 aa linear PRI 23-MAY-1995  
DEFINITION cytotoxic T-lymphocyte-associated serine esterase 1 [Homo sapiens].  
ACCESSION AAA67124  
VERSION AAA67124.1 GI:306682  
DBSOURCE locus HUMCTLA1 accession M38193.1  
KEYWORDS .  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 247)  
AUTHORS Caputo, A., Sauer, D.E. and Rowe, P.B.  
TITLE Nucleotide sequence and genomic organization of a human T  
lymphocyte serine protease gene  
JOURNAL J. Immunol. 145 (2), 737-744 (1990)  
PUBMED 2365998  
COMMENT Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..247  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="14q11.2"  
Protein 1..247  
/product="cytotoxic T-lymphocyte-associated serine  
esterase 1"  
Region 21..243  
/region\_name="Tryp\_SPC"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cd00190"  
/db\_xref="CDD:29152"  
Site 21  
/site\_type="cleavage"  
/db\_xref="CDD:29152"  
Site order(64,108,203)  
/site\_type="active"  
/db\_xref="CDD:29152"  
Site order(197,218,220)  
/site\_type="other"  
/note="substrate binding sites"  
/db\_xref="CDD:29152"  
CDS 1..247  
/gene="CTLA1"  
/coded\_by="join(M38193.1:910..964,M38193.1:2008..2155,  
M38193.1:2611..2746,M38193.1:2952..3212,  
M38193.1:3856..3999)"  
/db\_xref="GDB:G00-120-744"

ORIGIN  
1 mqpilllllaf lllpradage iiggheakph srpymaylmi wdqkslkrcg gfliqddfv1  
61 taahcwgsi nvtlgahnik eqeptqqfip vkrpiphpay npknfsndim llqlerkakr  
121 travqplrlp snkagvkpgq tcsvagwgqt aplgkhshtl qevkmtvqed rkcesdlrhy  
181 ydstielcvg dpeikktsfk gdsggplvcn kvaggivsyg rnngmpprac tkvssfvhwi  
241 kktmkry  
//



**P04187 (GRAB\_MOUSE) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified March 8, 2011. Version 119.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Granzyme B(G,H)</b> EC=3.4.21.79 <i>Alternative name(s):</i> CTLA-1 Cytotoxic cell protease 1 Short name=CCP1 Fragmentin-2
Gene names	Name: <b>Gzmb</b> Synonyms: Ctla-1, Ctla1
Organism	<b>Mus musculus (Mouse)</b>
Taxonomic identifier	10090 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires › Rodentia › Sciurognathi › Muroidea › Muridae › Murinae › Mus › Mus

**Protein attributes**

Sequence length	247 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	This enzyme is necessary for target cell lysis in cell-mediated immune responses. It cleaves after Asp. Seems to be linked to an activation cascade of caspases (aspartate-specific cysteine proteases) responsible for apoptosis execution. Cleaves caspase-3, -7, -9 and 10 to give rise to active enzymes mediating apoptosis <span>By similarity</span> .
Catalytic activity	Preferential cleavage: -Asp- -Xaa- >> -Asn- -Xaa- > -Met- -Xaa-, -Ser- -Xaa-.
Subcellular location	Cytoplasmic granule. Note: Cytoplasmic granules of cytolytic T-lymphocytes and natural killer cells.
Sequence similarities	Belongs to the peptidase S1 family. Granzyme subfamily. Contains 1 peptidase S1 domain.

Ontologies

Keywords




Biological process	Apoptosis Cytolysis
Domain	Signal
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Glycoprotein Zymogen
Technical term	3D-structure Direct protein sequencing

Gene Ontology (GO)

Biological process	T cell mediated cytotoxicity Inferred from mutant phenotype. Source: MGI  apoptosis Inferred from electronic annotation. Source: UniProtKB-KW  cytolysis Inferred from electronic annotation. Source: UniProtKB-KW  induction of apoptosis by granzyme Inferred from direct assay. Source: MGI  proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	cytoplasm Inferred from direct assay. Source: MGI
Molecular function	protein binding Inferred from physical interaction. Source: MGI  serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro

Complete GO annotation...

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
Molecule processing					
	Signal peptide	1 – 18	18		
	Propeptide	19 – 20	2	Activation peptide	PRO_0000027401
	Chain	21 – 247	227	Granzyme B(G,H)	PRO_0000027402

Regions

<input checked="" type="checkbox"/>	Domain	21 – 245	225	Peptidase S1		
-------------------------------------	--------	----------	-----	--------------	--	--

Sites

<input type="checkbox"/>	Active site	64	1	Charge relay system By similarity		
<input type="checkbox"/>	Active site	108	1	Charge relay system By similarity		
<input type="checkbox"/>	Active site	203	1	Charge relay system By similarity		

Amino acid modifications

<input type="checkbox"/>	Glycosylation	71	1	N-linked (GlcNAc...) Potential		
<input type="checkbox"/>	Glycosylation	182	1	N-linked (GlcNAc...) Potential		
<input type="checkbox"/>	Disulfide bond	49 ↔ 65		By similarity		
<input type="checkbox"/>	Disulfide bond	142 ↔ 209		By similarity		
<input type="checkbox"/>	Disulfide bond	173 ↔ 188		By similarity		

Secondary structure

1 ..... 247

■ Helix ■ Strand ■ Turn

Details...

Sequences

Sequence	Length	Mass (Da)
<input checked="" type="checkbox"/> P04187 [UniParc]. Last modified March 20, 1987. Version 1. Checksum: 996BCD199965C6D6	FASTA 247	27,470
<div><div>102030405060</div><div>MKILLLLLTLSLASRTKAGEIIGGHEVKPHSRPYMALLSIKDQQPEAICG GFLIREDFVL</div><div>708090100110120</div><div>TAAHCEGSII NVTLGAHNIK EQEKTQQVIP MVKCIPHPDY NPKTFSNDIM LLKLKSKAKR</div></div>		

```
      130      140      150      160      170      180
TRAVRPLNLP RRNVNVKPGD VCYVAGWGRM APMGKYSNTL QEVELTVQKD RECESYFKNR

      190      200      210      220      230      240
YNKTNQICAG DPKTKRASFR GDSGGPLVCK KVAAGIVSYG YKDGSPPRAF TKVSSFLSWI

      KKTMKSS
```




« Hide

## References

	« Hide 'large scale' references
[1]	<b>"Novel serine proteases encoded by two cytotoxic T lymphocyte-specific genes."</b> Lobe C.G., Finlay B.B., Paranchych W., Paetkau V.H., Bleackley R.C. Science 232:858-861(1986) [PubMed: 3518058] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].
[2]	<b>"Organization of two genes encoding cytotoxic T lymphocyte-specific serine proteases CCPI and CCPII."</b> Lobe C.G., Upton C., Duggan B., Ehrman N., Letellier M., Bell J., McFadden G., Bleackley R.C. Biochemistry 27:6941-6946(1988) [PubMed: 3264185] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA].
[3]	<b>"The inducible cytotoxic T-lymphocyte-associated gene transcript CTLA-1 sequence and gene localization to mouse chromosome 14."</b> Brunet J.-F., Dosseto M., Denizot F., Mattei M.-G., Clark W.R., Haqqi T.M., Ferrier P., Nabholz M., Schmitt-Verhulst A.-M., Luciani M.-F., Golstein P. Nature 322:268-271(1986) [PubMed: 3090449] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].
[4]	<b>"The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC)."</b> The MGC Project Team Genome Res. 14:2121-2127(2004) [PubMed: 15489334] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA]. <u>Strain:</u> FVB/N. <u>Tissue:</u> Mammary gland.
[5]	<b>"Genetic mapping of 40 cDNA clones on the mouse genome by PCR."</b> Ko M.S., Wang X., Horton J.H., Hagen M.D., Takahashi N., Maezaki Y., Nadeau J.H. Mamm. Genome 5:349-355(1994) [PubMed: 8043949] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE OF 227-247. <u>Strain:</u> C57BL/6J.
[6]	<b>"A family of serine esterases in lytic granules of cytolytic T lymphocytes."</b> Masson D., Tschopp J. Cell 49:679-685(1987) [PubMed: 3555842] [Abstract] <u>Cited for:</u> PROTEIN SEQUENCE OF 21-40.
[7]	<b>"Comparative molecular model building of two serine proteinases from cytotoxic T lymphocytes."</b> Murphy M.E.P., Moulton J., Bleackley R.C., Gershenson H., Weissman I.L., James M.N.G. Proteins 4:190-204(1988) [PubMed: 3237717] [Abstract] <u>Cited for:</u> 3D-STRUCTURE MODELING.
+	Additional computationally mapped references.



**Cross-references****Sequence databases**

 EMBL	X04072 mRNA. Translation: CAA27715.1.
 GenBank	M12302 mRNA. Translation: AAA37383.1.
 DDBJ	M22526 Genomic DNA. Translation: AAB61756.1.
	BC002085 mRNA. Translation: AAH02085.1.
	U05707 Genomic DNA. Translation: AAB60470.1.




IPI	IPI00114066.
-----	--------------

PIR	PRMSCL. A94288.
-----	-----------------

RefSeq	NP_038570.1. NM_013542.2.
--------	---------------------------

UniGene	Mm.14874.
---------	-----------

**3D structure databases**

 PDB	Entry	Method	Resolution	Chain	Positions	PDBsum
 RCSB PDB			(Å)			
 PDBj	2CP1	model	-	A	21-247	[»]

ProteinModelPortal	P04187.
--------------------	---------

SMR	P04187. Positions 21-245.
-----	---------------------------

ModBase	Search...
---------	-----------

**Protein-protein interaction databases**

DIP	DIP-562N.
-----	-----------

STRING	P04187.
--------	---------

**Protein family/group databases**

MEROPS	S01.136.
--------	----------

**Proteomic databases**

PRIDE	P04187.
-------	---------

**Genome annotation databases**

Ensembl	ENSMUST00000015581; ENSMUSP00000015581; ENSMUSG00000015437.
---------	--

GeneID	14939.
--------	--------

KEGG	mmu:14939.
------	------------

UCSC	uc007ubv.1. mouse.
------	--------------------

**Organism-specific databases**

CTD	14939.
-----	--------

MGI	MGI:109267. Gzmb.
-----	-------------------

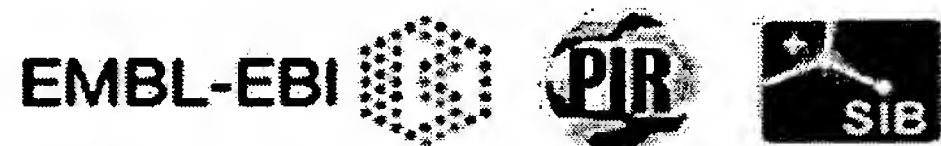
**Phylogenomic databases**

eggNOG	roNOG09047.
--------	-------------

HOGENOM	HBG755338.
HOVERGEN	HBG013304.
InParanoid	P04187.
OMA	YDSTVEL.
OrthoDB	EOG4KWJTW.
PhylomeDB	P04187.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.21.79. 244.
<b>Gene expression databases</b>	
ArrayExpress	P04187.
Bgee	P04187.
CleanEx	MM_GZMB.
Genevestigator	P04187.
GermOnline	ENSMUSG00000015437. Mus musculus.
<b>Family and domain databases</b>	
InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
NextBio	287263.
SOURCE	Search...
<b>Entry information</b>	
Entry name	GRAB_MOUSE
Accession	Primary (citable) accession number: <b>P04187</b>
Entry history	Integrated into March 20, 1987 UniProtKB/Swiss- Prot:

	Last sequence update: March 20, 1987 Last modified: March 8, 2011 This is version 119 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Relevant documents</b>	
MGD cross-references Mouse Genome Database (MGD) cross-references in UniProtKB/Swiss-Prot	
PDB cross-references Index of Protein Data Bank (PDB) cross-references	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



**P51124 (GRAM\_HUMAN) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 96.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Granzyme M</b> EC=3.4.21.- <i>Alternative name(s):</i> Met-1 serine protease Short name=Hu-Met-1 Met-ase Natural killer cell granular protease
Gene names	Name: <b>GZMM</b> Synonyms: MET1
Organism	<b>Homo sapiens (Human)</b> [Complete proteome]
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo

**Protein attributes**

Sequence length	257 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Cleaves peptide substrates after methionine, leucine, and norleucine. Physiological substrates include EZR, alpha-tubulins and the apoptosis inhibitor BIRC5/Survivin. Promotes caspase activation and subsequent apoptosis of target cells.
Subcellular location	Secreted. Cytoplasmic granule. Note: Granules of large granular lymphocytes.
Tissue specificity	Highly and constitutively expressed in activated natural killer (NK) cells.
Sequence similarities	Belongs to the peptidase S1 family. Granzyme subfamily. Contains 1 peptidase S1 domain.

**Ontologies****Keywords**



Biological process	Apoptosis Cytolysis Immunity Innate immunity
Cellular component	Secreted
Coding sequence diversity	Polymorphism
Domain	Signal
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Glycoprotein Zymogen
Technical term	3D-structure Complete proteome
<b>Gene Ontology (GO)</b>	
Biological process	apoptosis Inferred from electronic annotation. Source: UniProtKB-KW  cytolysis Inferred from electronic annotation. Source: UniProtKB-KW  innate immune response Inferred from electronic annotation. Source: UniProtKB-KW  proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	extracellular region Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro
Complete GO annotation...	

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature id
<b>Molecule processing</b>					
<input checked="" type="checkbox"/> Signal peptide	1 – 23	23	Potential		
<input type="checkbox"/> Propeptide	24 – 25	2	Activation peptide Potential		PRO_00000
<input checked="" type="checkbox"/> Chain	26 – 257	232	Granzyme M		PRO_00000
<b>Regions</b>					
<input checked="" type="checkbox"/> Domain	26 – 254	229	Peptidase S1		

Sites

<input type="checkbox"/>	Active site	66	1	Charge relay system		
<input type="checkbox"/>	Active site	111	1	Charge relay system		
<input type="checkbox"/>	Active site	207	1	Charge relay system		

Amino acid modifications

<input type="checkbox"/>	Glycosylation	177	1	N-linked (GlcNAc...) Potential		
<input type="checkbox"/>	Disulfide bond	51 ↔ 67				
<input type="checkbox"/>	Disulfide bond	145 ↔ 213				
<input type="checkbox"/>	Disulfide bond	176 ↔ 192				
<input type="checkbox"/>	Disulfide bond	203 ↔ 230				

Natural variations

<input type="checkbox"/>	Natural variant	221	1	R → G. [dbSNP:rs1599882] Ref.1 Ref.2 Ref.4		VAR_05182
--------------------------	-----------------	-----	---	--	--	-----------

Secondary structure

1 . . . . .

■ Helix ■ Strand ■ Turn

Details...

Sequences

Sequence	Length	Mass (Da)
<div><div><input type="checkbox"/> P51124 [UniParc].</div><div>FASTA</div><div>257</div><div>27,545</div></div> <div>Last modified May 18, 2010. Version 2. Checksum: B4E815CE455F7371</div> <div><div>102030405060</div><div>MEACVSSLLV LALGALSVGS SFGTQIIIGR EVIPHSRPYM ASLQRNGSHL CGGVLVHPKW</div><div>708090100110120</div><div>VLTAAHCLAQ RMAQLRLVLG LHTLDSPGLT FHIKAAIQHP RYKPVPALEN DLALLQLDGK</div><div>130140150160170180</div><div>VKPSRTIRPL ALPSKRQVVA AGTRCSMAGW GLTHQGGRLS RVLRELDLQV LDTRMCNNSR</div><div>190200210220230240</div><div>FWNGSLSPSM VCLAADSKDQ APCKGDSGGP LVC GKGRVLA RVLSFSSRVC TDIFKPPVAT</div><div>250</div><div>AVAPYVSWIR KVTGRSA</div></div>		

« Hide

## References

	« Hide 'large scale' references
[1]	<p><b>"Met-ase: cloning and distinct chromosomal location of a serine protease preferentially expressed in human natural killer cells."</b></p> <p>Smyth M.J., Sayers T.J., Wiltout T., Powers J.C., Trapani J.A.  J. Immunol. 151:6195-6205(1993) [PubMed: 8245461] [Abstract]  Cited for: NUCLEOTIDE SEQUENCE [MRNA], VARIANT GLY-221.</p>
[2]	<p><b>"The human Met-ase gene (GZMM): structure, sequence, and close physical linkage to the serine protease gene cluster on 19p13.3."</b></p> <p>Pilat D., Fink T.M., Obermaier-Skrobanek B., Zimmer M., Wekerle H., Lichter P., Jenne D.E.  Genomics 24:445-450(1994) [PubMed: 7713495] [Abstract]  Cited for: NUCLEOTIDE SEQUENCE [GENOMIC DNA], VARIANT GLY-221.</p>
[3]	<p><b>"The DNA sequence and biology of human chromosome 19."</b></p> <p>Grimwood J., Gordon L.A., Olsen A.S., Terry A., Schmutz J., Lamerdin J.E., Hellsten U., Goodstein D., Couronne O., Tran-Gyamfi M., Aerts A., Altherr M., Ashworth L., Bajorek E., Black S., Branscomb E., Caenepeel S., Carrano A.V., Lucas S.M.  Nature 428:529-535(2004) [PubMed: 15057824] [Abstract]  Cited for: NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].</p>
[4]	<p><b>"The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC)."</b></p> <p>The MGC Project Team  Genome Res. 14:2121-2127(2004) [PubMed: 15489334] [Abstract]  Cited for: NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], VARIANT GLY-221.  Tissue: Pancreas and Spleen.</p>
[5]	<p><b>"NK cell protease granzyme M targets alpha-tubulin and disorganizes the microtubule network."</b></p> <p>Bovenschen N., de Koning P.J., Quadir R., Broekhuizen R., Damen J.M., Froelich C.J., Slijper M., Kummer J.A.  J. Immunol. 180:8184-8191(2008) [PubMed: 18523284] [Abstract]  Cited for: FUNCTION, TISSUE SPECIFICITY, SUBSTRATES.</p>
[6]	<p><b>"Cleavage of survivin by Granzyme M triggers degradation of the survivin-X-linked inhibitor of apoptosis protein (XIAP) complex to free caspase activity leading to cytolysis of target tumor cells."</b></p> <p>Hu D., Liu S., Shi L., Li C., Wu L., Fan Z.  J. Biol. Chem. 285:18326-18335(2010) [PubMed: 20406824] [Abstract]  Cited for: FUNCTION.</p>
[7]	<p><b>"Structural basis for proteolytic specificity of the human apoptosis-inducing granzyme M."</b></p> <p>Wu L., Wang L., Hua G., Liu K., Yang X., Zhai Y., Bartlam M., Sun F., Fan Z.  J. Immunol. 183:421-429(2009) [PubMed: 19542453] [Abstract]  Cited for: X-RAY CRYSTALLOGRAPHY (1.96 ANGSTROMS) OF 26-257, ACTIVE SITE, DISULFIDE BONDS.</p>
+	Additional computationally mapped references.

**Cross-references****Sequence databases**

<input type="radio"/> EMBL <input type="radio"/> GenBank <input type="radio"/> DDBJ	L23134 mRNA. Translation: AAA59582.1. L36922 Genomic DNA. Translation: AAA57262.1. L36936 Genomic DNA. Translation: AAA57257.1. AC011556 Genomic DNA. No translation available. BC025701 mRNA. Translation: AAH25701.1.
---	---

IPI	IPI00016282.
PIR	A55634.
RefSeq	NP_005308.1. NM_005317.2.
UniGene	Hs.465511.

**3D structure databases**

<input type="radio"/> PDBe <input type="radio"/> RCSB PDB <input type="radio"/> PDBj	Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
	1LZP	model	-	A	26-257	[»]
	2ZGC	X-ray	1.96	A	26-257	[»]
	2ZGH	X-ray	2.17	A	26-257	[»]
	2ZGJ	X-ray	2.30	A	26-257	[»]
	2ZKS	X-ray	2.70	A	26-257	[»]

ProteinModelPortal	P51124.
SMR	P51124. Positions 26-256.
ModBase	Search...

**Protein-protein interaction databases**

STRING	P51124.
--------	---------

**Protein family/group databases**

MEROPS	S01.139.
--------	----------

**Proteomic databases**

PRIDE	P51124.
-------	---------

**Genome annotation databases**

Ensembl	ENST00000264553; ENSP00000264553; ENSG00000197540.
GeneID	3004.
KEGG	hsa:3004.
UCSC	uc002low.1. human.

**Organism-specific databases**

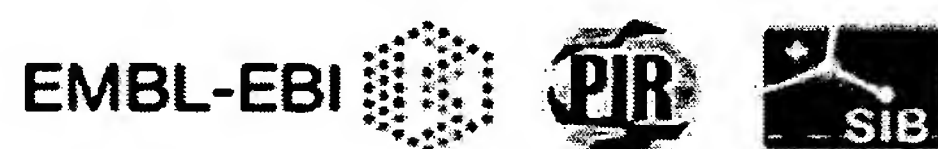
CTD	3004.
GeneCards	GC19P000314.
H-InvDB	HIX0202841.
HGNC	HGNC:4712. GZMM.



MIM	600311. gene.
neXtProt	NX_P51124.
PharmGKB	PA29090.
GenAtlas	Search...
<b>Phylogenomic databases</b>	
eggNOG	prNOG21107.
GeneTree	ENSGT00580000081367.
HOGENOM	HBG755338.
HOVERGEN	HBG013304.
InParanoid	P51124.
OrthoDB	EOG4TMR2V.
<b>Gene expression databases</b>	
ArrayExpress	P51124.
Bgee	P51124.
CleanEx	HS_GZMM.
Genevestigator	P51124.
GermOnline	ENSG00000197540. Homo sapiens.
<b>Family and domain databases</b>	
InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
NextBio	11912.
SOURCE	Search...
<b>Entry information</b>	
Entry name	GRAM_HUMAN
Accession	Primary (citable) accession number: <b>P51124</b>

Entry history	Integrated into UniProtKB/Swiss-Prot: October 1, 1996 Last sequence update: May 18, 2010 Last modified: April 5, 2011 This is version 96 of the entry and version 2 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
Disclaimer	Any medical or genetic information present in this entry is provided for research, educational and informational purposes only. It is not in any way intended to be used as a substitute for professional medical advice, diagnosis, treatment or care.
<b>Relevant documents</b>	
Human chromosome 19 Human chromosome 19: entries, gene names and cross-references to MIM	
Human entries with polymorphisms or disease mutations List of human entries with polymorphisms or disease mutations	
Human polymorphisms and disease mutations Index of human polymorphisms and disease mutations	
MIM cross-references Online Mendelian Inheritance in Man (MIM) cross-references in UniProtKB/Swiss-Prot	
PDB cross-references Index of Protein Data Bank (PDB) cross-references	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



## Protein

Translations of Life

Display Settings: GenPept

## metase [Homo sapiens]

GenBank: AAA59582.1

[FASTA](#) [Graphics](#)Go to:

LOCUS AAA59582 257 aa linear PRI 07-JAN-1995  
DEFINITION metase [Homo sapiens].  
ACCESSION AAA59582  
VERSION AAA59582.1 GI:508988  
DBSOURCE locus HUMMET1A accession L23134.1  
KEYWORDS .  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 257)  
AUTHORS Smyth,M.J., Sayers,T.J., Wilttrout,T., Powers,J.C. and Trapani,J.A.  
TITLE Met-ase: cloning and distinct chromosomal location of a serine  
protease preferentially expressed in human natural killer cells  
J. Immunol. 151 (11), 6195-6205 (1993)  
PUBMED 8245461  
COMMENT On Jul 13, 1994 this sequence version replaced gi:438641.  
Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..257  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_line="Human Lopez natural killer cell leukemia"  
Protein 1..257  
/product="metase"  
sig\_peptide 1..25  
/gene="MET-1"  
mat\_peptide 26..257  
/gene="MET-1"  
/product="metase"  
Region 26..252  
/region\_name="Tryp\_SpC"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cd00190"  
/db\_xref="CDD:29152"  
Site 26  
/site\_type="cleavage"  
/db\_xref="CDD:29152"  
Site order(66,111,207)  
/site\_type="active"  
/db\_xref="CDD:29152"  
Site order(201,224,226)  
/site\_type="other"  
/note="substrate binding sites"  
/db\_xref="CDD:29152"  
CDS 1..257  
/gene="MET-1"  
/coded\_by="join(L23134.1:1..75,L23134.1:76..771,  
L23134.1:785..787)"  
ORIGIN  
1 meacvssllv lalgalsvgs sfgtqiiggr eviphsrpym aslqrngshl cggvlvhpkw  
61 vltaahclaq rmaqlrlvlq lhtldspglt fhikaaighp rykvpvalen dlallqldgk  
121 vkpsertrpl alpskrqva agtrcsmagw glthqggrls rvlreldlqv ldtrmcnnsr  
181 fwngslspsm vclaadskdq apckgdsggp lvcgkgrvla gvlsfssrvc tdfkppvat  
241 avapyvswir kvtrgsa  
//

## Protein

Translations of Life

Display Settings: GenPept

## portion of prepropeptide [Homo sapiens]

GenBank: AAA57262.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAA57262 18 aa linear PRI 15-DEC-1994  
DEFINITION portion of prepropeptide [Homo sapiens].  
ACCESSION AAA57262  
VERSION AAA57262.1 GI:602469  
DBSOURCE locus HUMPOP accession [L36922.1](#)  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM [Homo sapiens](#)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 18)  
AUTHORS Smyth,M.J., Sayers,T.J., Wiltout,T., Powers,J.C. and Trapani,J.A.  
TITLE Met-ase: cloning and distinct chromosomal location of a serine  
protease preferentially expressed in human natural killer cells  
J. Immunol. 151 (11), 6195-6205 (1993)  
JOURNAL  
PUBMED [8245461](#)  
REFERENCE 2 (residues 1 to 18)  
AUTHORS Pilat,D., Fink,T., Obermaier-Skrobanek,B., Zimmer,M., Wekerle,H.,  
Lichter,P. and Jenne,D.E.  
TITLE The human Met-ase gene (GZMM): structure, sequence, and close  
physical linkage to the serine protease gene cluster on 19p13.3  
JOURNAL Genomics 24 (3), 445-450 (1994)  
PUBMED [7713495](#)  
COMMENT Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..18  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
Protein 1..18  
/name="portion of prepropeptide"  
CDS 1..18  
/coded\_by="L36922.1:845..899"  
ORIGIN  
1 meacvssllv lalgalsv  
//



**Q03238 (GRAM\_RAT) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 83.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Granzyme M</b> EC=3.4.21.- <i>Alternative name(s):</i> Met-ase Natural killer cell granular protease RNK-Met-1
Gene names	Name: <b>Gzmm</b>
Organism	<b>Rattus norvegicus (Rat)</b>
Taxonomic identifier	10116 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires › Rodentia › Sciurognathi › Muroidea › Muridae › Murinae › Rattus

**Protein attributes**

Sequence length	258 AA.
Sequence status	Fragment.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Cleaves peptide substrates after methionine, leucine, and norleucine. Physiological substrates include EZR, alpha-tubulins and the apoptosis inhibitor BIRC5/Survivin. Promotes caspase activation and subsequent apoptosis of target cells <span>By similarity</span>
Subcellular location	Secreted. Cytoplasmic granule. Note: Granules of large granular lymphocytes.
Sequence similarities	Belongs to the peptidase S1 family. Granzyme subfamily. Contains 1 peptidase S1 domain.

**Ontologies**

<b>Keywords</b>	
Biological process	Apoptosis Cytolysis

	Immunity Innate immunity
Cellular component	Secreted
Domain	Signal
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Glycoprotein Zymogen
Technical term	Direct protein sequencing
<b>Gene Ontology (GO)</b>	
Biological process	apoptosis Inferred from electronic annotation. Source: UniProtKB-KW cytolysis Inferred from electronic annotation. Source: UniProtKB-KW innate immune response Inferred from electronic annotation. Source: UniProtKB-KW proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	extracellular region Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro

Complete GO annotation...

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
<b>Molecule processing</b>					
<input type="checkbox"/> Signal peptide	1 – ?		Potential		
<input type="checkbox"/> Propeptide	? – 20		Activation peptide		PRO_0000027
<input checked="" type="checkbox"/> Chain	21 – 258	238	Granzyme M		PRO_0000027
<b>Regions</b>					
<input checked="" type="checkbox"/> Domain	21 – 250	230	Peptidase S1		

Sites					
<input type="checkbox"/>	Active site	61	1	Charge relay system By similarity	
<input type="checkbox"/>	Active site	107	1	Charge relay system By similarity	
<input type="checkbox"/>	Active site	204	1	Charge relay system By similarity	
Amino acid modifications					
<input type="checkbox"/>	Glycosylation	174	1	N-linked (GlcNAc...) Potential	
<input type="checkbox"/>	Glycosylation	225	1	N-linked (GlcNAc...) Potential	
<input type="checkbox"/>	Disulfide bond	46 ↔ 62		By similarity	
<input type="checkbox"/>	Disulfide bond	142 ↔ 210		By similarity	
<input type="checkbox"/>	Disulfide bond	173 ↔ 189		By similarity	
<input type="checkbox"/>	Disulfide bond	200 ↔ 226		By similarity	
Experimental info					
<input type="checkbox"/>	Non-terminal residue	1	1		

Sequences

Sequence	Length	Mass (Da)
<div><div><div><div><input type="checkbox"/></div><div>Q03238 [UniParc].</div><div>Last modified June 1, 1994. Version 1.</div><div>Checksum: B89DC10EF54DF495</div></div><div><div>102030405060</div><div>LLLLLALKTLWAVGNRFEAQIIGGREAVPHSRPYMVSLQNTKSHMCGGVLVHQKWVLTAA</div><div>708090100110120</div><div>HCLSEPLQQCLKVFGHLHSLHDPQDPGLTFYIKQAIKHPGYNLKYENDLALLKLDGRVKPS</div><div>130140150160170180</div><div>KNVKPLALPRKPRDKPAEGSRCSTAGWGITHQRGQLAKSLQELDLRLLDTRMCNNSRFWN</div><div>190200210220230240</div><div>GVLTDSDLCLKAGAKGQAPCKGDSGGPLVCGKGKVDGILSFSSKNCTDIFKPTVATAVAP</div><div>250</div><div>YSSWIRKVIGRWSPQPLT</div></div></div></div>	FASTA	25828,339

« Hide

## References

- [1] **"Purification and cloning of a novel serine protease, RNK-Met-1, from the granules of a rat natural killer cell leukemia."**  
 Smyth M.J., Wiltout T., Trapani J.A., Ottaway K.S., Sowder R., Henderson L.E., Kam C.-M., Powers J.C., Young H.A., Sayers T.J.  
 J. Biol. Chem. 267:24418-24425(1992) [PubMed: 1447189] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA], PROTEIN SEQUENCE OF 21-44.
- + Additional computationally mapped references.

## Cross-references

### Sequence databases

<input checked="" type="radio"/> EMBL <input checked="" type="radio"/> GenBank <input checked="" type="radio"/> DDBJ	L05175 mRNA. Translation: AAA42056.1.
--	---------------------------------------

IPI	IPI00365967.
-----	--------------

PIR	A45161.
-----	---------

UniGene	Rn.9838.
---------	----------

### 3D structure databases

ProteinModelPortal	Q03238.
--------------------	---------

ModBase	Search...
---------	-----------

### Protein-protein interaction databases

STRING	Q03238.
--------	---------

### Protein family/group databases

MEROPS	S01.139.
--------	----------

### Genome annotation databases

Ensembl	ENSRNOT00000011086; ENSRNOP00000011086; ENSRNOG00000030530.
---------	--

### Organism-specific databases

RGD	620022. Gzmm.
-----	---------------

### Phylogenomic databases

eggNOG	maNOG18874.
--------	-------------

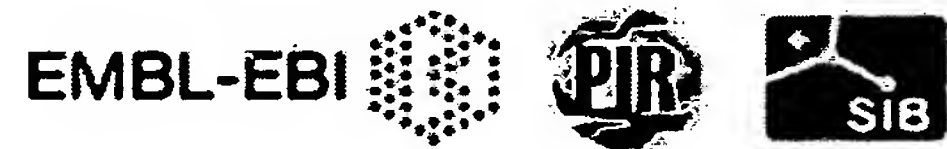
GeneTree	ENSGT00580000081367.
----------	----------------------

HOVERGEN	HBG013304.
----------	------------



InParanoid	Q03238.
OrthoDB	EOG4TMR2V.
<b>Gene expression databases</b>	
ArrayExpress	Q03238.
Genevestigator	Q03238.
GermOnline	ENSRNOG00000030530. Rattus norvegicus.
<b>Family and domain databases</b>	
InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Entry information</b>	
Entry name	GRAM_RAT
Accession	Primary (citable) accession number: <b>Q03238</b>
Entry history	Integrated into June 1, 1994 UniProtKB/Swiss-Prot: Last sequence update: June 1, 1994 Last modified: April 5, 2011 This is version 83 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Relevant documents</b>	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



**O08643** (O08643\_MOUSE) ★ Unreviewed, UniProtKB/TrEMBL

Last modified March 8, 2011. Version 80.

**Names and origin**

Protein names	<i>Submitted name:</i> Granzyme M <a href="#">EMBL BAA78336.1</a> <i>Submitted name:</i> Granzyme M (Lymphocyte met-ase 1) <a href="#">EMBL AAI39218.1</a> <i>Submitted name:</i> Granzyme M (Lymphocyte met-ase 1), isoform CRA_b <a href="#">EMBL EDL31677.1</a> <i>Submitted name:</i> Granzyme Met-ase-1 precursor <a href="#">EMBL AAB51606.1</a>
Gene names	Name: <b>Gzmm</b> <a href="#">MGI 99549</a> Synonyms: <b>LMet-1</b> <a href="#">EMBL AAB51606.1</a> , <b>MMET-1</b> <a href="#">EMBL BAA78336.1</a> ORF <b>mCG_13391</b> <a href="#">EMBL EDL31677.1</a> Names:
Organism	<b>Mus musculus (Mouse)</b> <a href="#">EMBL AAB51606.1</a>
Taxonomic identifier	10090 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires › Rodentia › Sciurognathi › Muroidea › Muridae › Murinae › Mus › Mus

**Protein attributes**

Sequence length	264 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at transcript level.

**General annotation (Comments)**

Sequence similarities	Belongs to the peptidase S1 family. <a href="#">RuleBase RU000360V4</a>
-----------------------	---



**Ontologies****Keywords**

Domain	Signal <a href="#">EMBL AAB51606.1</a>
Molecular function	Hydrolase Protease Serine protease <a href="#">RuleBase RU004260V0</a>


**Gene Ontology (GO)**

Biological process	cytolysis Inferred from direct assay. Source: MGI  proteolysis Inferred from electronic annotation. Source: InterPro
Molecular function	protein binding Inferred from physical interaction. Source: MGI  serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro
Complete GO annotation...	

Sequence annotation (Features)

Feature key	Position (s)	Length	Description	Graphical view	Feature identifier
Molecule processing					
	Signal peptide	1 – 26	26	<div>Potential</div> <div>EMBL AAB51606.1</div>	
	Chain	27 – 264	238	<div>Potential</div> <div>EMBL AAB51606.1</div>	PRO_5000142980

Sequences

Sequence	Length	Mass (Da)
<div> O08643 [UniParc]. Last modified July 1, 1997. Version 1. Checksum: 496E644D2630CD14</div>	FASTA 264	29,122
<div><div><div>102030405060</div><div>MEVCWSLLLLL LALKTLWAAG NRFETQIIGG REAVPHSRPY MASLQKAKSH VCGGVLVHRK</div></div><div><div>708090100110120</div><div>WVLTAHCLS EPLQNLKLVL GLHNLHDLQD PGLTFYIREA IKHPGYNHKY ENDLALLKLD</div></div><div><div>130140150160170180</div><div>RRVQPSKNVK PLALPRKPRS KPAEGTWCST AGWGMTHQGG PRARALQELD LRVLDTQMCN</div></div><div><div>190200210220230240</div><div>NSRFWNGVLI DSMLCLKAGS KSQAPCKGDS GGPLVCGKGQ VDGILSFSSK TCTDIFKPPV</div></div><div><div>250260</div><div>ATAVAPYSSW IRKVIGRWSP QSLV</div></div></div> <div>« Hide</div>		

References

	« Hide 'large scale' references
[1]	"Cloning and expression of the recombinant mouse natural killer cell granzyme Met-ase-1." Kelly J.M., O'Connor M.D., Hulett M.D., Thia K.Y.T., Smyth M.J.



	Immunogenetics 44:340-350(1996) [PubMed: 8781119] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE. <u>Strain:</u> BALB/c <a href="#">EMBL AAB51606.1</a> <u>Tissue:</u> Spleen <a href="#">EMBL AAB51606.1</a>
[2]	<b>"A comparison of whole-genome shotgun-derived mouse chromosome 16 and the human genome."</b> Mural R.J., Adams M.D., Myers E.W., Smith H.O., Miklos G.L., Wides R., Halpern A., Li P.W., Sutton G.G., Nadeau J., Salzberg S.L., Holt R.A., Kodira C.D., Lu F., Chen L., Deng Z., Evangelista C.C., Gan W. <a href="#">↔</a> Stephenson L.D. Science 296:1661-1671(2002) [PubMed: 12040188] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE. <u>Strain:</u> Mixed <a href="#">EMBL EDL31677.1</a>
[3]	Mural R.J., Adams M.D., Myers E.W., Smith H.O., Venter J.C. Submitted (SEP-2005) to the EMBL/GenBank/DDBJ databases <u>Cited for:</u> NUCLEOTIDE SEQUENCE. <u>Strain:</u> Mixed <a href="#">EMBL EDL31677.1</a>
[4]	<b>"The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC)."</b> The MGC Project Team Genome Res. 14:2121-2127(2004) [PubMed: 15489334] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA]. <u>Tissue:</u> Brain <a href="#">EMBL AAI39218.1</a>
[5]	<b>"High expression of alternative transcript of granzyme M in the mouse retina."</b> Taniguchi M., Tani N., Suemoto T., Ishimoto I., Shiosaka S., Yoshida S. Neurosci. Res. 34:115-123(1999) [PubMed: 10498337] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE.
+	Additional computationally mapped references.

## Cross-references

### Sequence databases

<input checked="" type="radio"/> EMBL	L76741 mRNA. Translation: AAB51606.1.
<input type="radio"/> GenBank	BC139217 mRNA. Translation: AAI39218.1.
<input type="radio"/> DDBJ	BC139218 mRNA. Translation: AAI39219.1.
	AB015728 mRNA. Translation: BAA78336.1.
	CH466553 Genomic DNA. Translation: EDL31677.1.
IPI	IPI00114906.
RefSeq	NP_032530.1. NM_008504.2.
UniGene	Mm.378960.

### 3D structure databases

HSSP	HSSP built from PDB template 1FI8 based on UniProtKB P18291.
ProteinModelPortal	O08643.
ModBase	Search...

### Protein-protein interaction databases

STRING	O08643.
--------	---------

**Protein family/group databases**

MEROPS	S01.139.
--------	----------

**Proteomic databases**

PRIDE	O08643.
-------	---------

**Genome annotation databases**

Ensembl	ENSMUST00000020549; ENSMUSP00000020549; ENSMUSG00000054206.
---------	--

GeneID	16904.
--------	--------

KEGG	mmu:16904.
------	------------

UCSC	uc007fzj.1. mouse.
------	--------------------

**Organism-specific databases**

CTD	16904.
-----	--------

MGI	MGI:99549. Gzmm.
-----	------------------

**Phylogenomic databases**

eggNOG	maNOG18874.
--------	-------------

HOGONOM	HBG755338.
---------	------------

HOVERGEN	HBG013304.
----------	------------

InParanoid	O08643.
------------	---------

OMA	KDQAPCK.
-----	----------

PhylomeDB	O08643.
-----------	---------

**Gene expression databases**

ArrayExpress	O08643.
--------------	---------

Bgee	O08643.
------	---------

Genevestigator	O08643.
----------------	---------

**Family and domain databases**

InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
----------	---

Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
------	--

PRINTS	PR00722. CHYMOTRYPSIN.
--------	------------------------

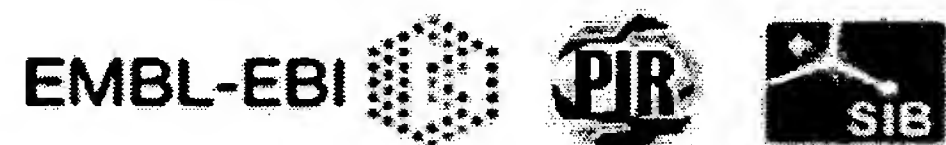
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
-------	---

SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
--------	--------------------------------

PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit.
---------	--

ProtoNet	PS00135. TRYPSIN_SER. 1 hit. [Graphical view] Search...
<b>Other Resources</b>	
NextBio	290930.
SOURCE	Search...
<b>Entry information</b>	
Entry name	O08643_MOUSE
Accession	Primary (citable) accession number: <b>O08643</b>
Entry history	Integrated into July 1, 1997 UniProtKB/TrEMBL: Last sequence July 1, 1997 update: Last modified: March 8, 2011 This is version 80 of the entry and version 1 of the sequence. [Complete history]
Entry status	Unreviewed (UniProtKB/TrEMBL)

© 2002–2011 UniProt Consortium | License &amp; Disclaimer | Contact



**P32197 (ELAS\_GADMO) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified March 2, 2010. Version 44.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Elastase</b> EC=3.4.21.-
Organism	<b>Gadus morhua (Atlantic cod)</b>
Taxonomic identifier	8049 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Actinopterygii › Neopterygii › Teleostei › Euteleostei › Neoteleostei › Acanthomorpha › Paracanthopterygii › Gadiformes › Gadidae › Gadus

**Protein attributes**

Sequence length	20 AA.
Sequence status	Fragment.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Digests most rapidly at the C-terminal side of alanine residues, but also cleaves at valine and leucine residues.
Sequence similarities	Belongs to the peptidase S1 family. Elastase subfamily. Contains 1 peptidase S1 domain.

**Ontologies****Keywords**



Molecular function	Hydrolase Protease Serine protease
Technical term	Direct protein sequencing

**Gene Ontology (GO)**


Molecular function	serine-type peptidase activity Inferred from electronic annotation. Source: UniProtKB-KW
Complete GO annotation...	



Sequence annotation (Features)

Feature key	Position (s)	Length	Description	Graphical view	Feature identifier
Molecule processing					
 Chain	1 – »20	»20	Elastase		PRO_000008867
Regions					
 Domain	1 – »20	»20	Peptidase S1		
Experimental info					
<input type="checkbox"/> Non-terminal residue	20	1			

Sequences

Sequence	Length	Mass (Da)
 P32197 [UniParc]. Last modified October 1, 1993. Version 1. Checksum: 51D4B08262AC84BC <div><div>1020</div><div>VVGGEVARAH SWPWQISLQY</div></div> « Hide	FASTA 20	2,284

References

[1]	"Properties of elastase from Atlantic cod, a cold-adapted proteinase." Asgeirsson B., Bjarnason J.B. Biochim. Biophys. Acta 1164:91-100(1993) [PubMed: 8518301] [Abstract] <u>Cited for:</u> PROTEIN SEQUENCE. <u>Tissue:</u> Intestine.
-----	--

Cross-references

Sequence databases	
PIR	S33787.
3D structure databases	
ModBase	Search...

**Protein family/group databases**

MEROPS	S01.153.
--------	----------

**Family and domain databases**

PROSITE	PS50240. TRYPSIN_DOM. Partial match. PS00134. TRYPSIN_HIS. Partial match. PS00135. TRYPSIN_SER. Partial match. [Graphical view]
---------	--

ProtoNet	Search...
----------	-----------

**Entry information**

Entry name	ELAS_GADMO
Accession	Primary (citable) accession number: <b>P32197</b>
Entry history	Integrated into      October 1, 1993 UniProtKB/Swiss- Prot: Last sequence      October 1, 1993 update: Last modified:      March 2, 2010 This is version 44 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program

**Relevant documents**

Peptidase families  
Classification of peptidase families and list of entries

SIMILARITY comments  
Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



**P23946 (CMA1\_HUMAN) ★ Reviewed, UniProtKB/Swiss-Prot**








Last modified April 5, 2011. Version 118.

Names and origin	
Protein names	<i>Recommended name:</i> <b>Chymase</b> EC=3.4.21.39 <i>Alternative name(s):</i> Alpha-chymase Mast cell protease I
Gene names	Name: <b>CMA1</b> Synonyms: CYH, CYM
Organism	<b>Homo sapiens (Human)</b> [Complete proteome]
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo
Protein attributes	
Sequence length	247 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.
General annotation (Comments)	
Function	Major secreted protease of mast cells with suspected roles in vasoactive peptide generation, extracellular matrix degradation, and regulation of gland secretion.
Catalytic activity	Preferential cleavage: Phe- -Xaa > Tyr- -Xaa > Trp- -Xaa > Leu- -Xaa.
Subcellular location	Secreted. Cytoplasmic granule. Note: Mast cell granules.
Tissue specificity	Mast cells in lung, heart, skin and placenta. Expressed in both normal skin and in urticaria pigmentosa lesions. <a href="#">Ref.3</a>
Sequence similarities	Belongs to the peptidase S1 family. Granzyme subfamily. Contains 1 peptidase S1 domain.
Ontologies	
<b>Keywords</b>	
Cellular component	Secreted
Coding sequence diversity	Polymorphism

Domain	Signal
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Glycoprotein Zymogen
Technical term	3D-structure Complete proteome Direct protein sequencing
<b>Gene Ontology (GO)</b>	
Biological process	interleukin-1 beta biosynthetic process Inferred from direct assay. Source: BHF-UCL  proteolysis Inferred from electronic annotation. Source: InterPro  regulation of inflammatory response Inferred by curator. Source: BHF-UCL
Cellular component	extracellular region Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	serine-type endopeptidase activity Non-traceable author statement <span>Ref.3</span> . Source: UniProtKB

Complete GO annotation...

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature
<b>Molecule processing</b>					
	Signal peptide	1 – 19	19		
	Propeptide	20 – 21	2	Activation peptide	PRO_01
	Chain	22 – 247	226	Chymase	PRO_01
<b>Regions</b>					
	Domain	22 – 245	224	Peptidase S1	
<b>Sites</b>					
	Active site	66	1	Charge relay system	
	Active site	110	1	Charge relay system	
	Active site	203	1	Charge relay system	



Amino acid modifications

<input type="checkbox"/>	Glycosylation	80	1	N-linked (GlcNAc...) Ref.9 Ref.1	I	
<input type="checkbox"/>	Glycosylation	103	1	N-linked (GlcNAc...) Ref.9	I	
<input type="checkbox"/>	Disulfide bond	51 ↔ 67			II	
<input type="checkbox"/>	Disulfide bond	144 ↔ 209			I I	
<input type="checkbox"/>	Disulfide bond	175 ↔ 188			II	

Natural variations

<input type="checkbox"/>	Natural variant	46	1	G → R. [dbSNP:rs5246]	I	VAR_0
<input type="checkbox"/>	Natural variant	66	1	H → R. [dbSNP:rs5247] Ref.6	I	VAR_0
<input type="checkbox"/>	Natural variant	98	1	R → H. [dbSNP:rs13306252]	I	VAR_0

Experimental info

<input type="checkbox"/>	Sequence conflict	28	1	C → S in AAB26828. Ref.7	I	
<input type="checkbox"/>	Sequence conflict	131 – 132	2	FP → AV AA sequence Ref.3	I	

Secondary structure

1. ....  
■ Helix ■ Strand ■ Turn

Details...



Sequences

Sequence	Length	Mass (Da)
<input checked="" type="checkbox"/> P23946 [UniParc]. Last modified March 1, 1992. Version 1. Checksum: DC1464A049ED6B00	FASTA 247	27,325
<div><div>102030405060</div><div>MLLLPLPLLL FLLCSRAEAG EIIGGTECKP HSRPYMAYLE IVTSNGPSKF CGGFLIRRNF</div><div>708090100110120</div><div>VLTAHCAGR SITVTLGAWN ITEEDTWQK LEVIKQFRHP KYNTSTLHHD IMLLKLKEKA</div><div>130140150160170180</div><div>SLTLAVGTLP FPSQNFVPP GRMCRVAGWG RTGVLKPGSD TLQEVKLRLM DPQACSHFRD</div><div>190200210220230240</div><div>FDHNLQLCVG NPRKTKSAFK GDSGGPLLCA GVAQGIVSYG RSDAKPPAVF TRISHYRPWI</div></div>		

NQILQAN

« Hide

## References

	« Hide 'large scale' references
[1]	<p><b>"Structure, chromosomal assignment, and deduced amino acid sequence of a human gene for mast cell chymase."</b>  Caughey G.H., Zerweck E.H., Vanderslice P.  J. Biol. Chem. 266:12956-12963(1991) [PubMed: 2071582] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA].</p>
[2]	<p><b>"Cloning of the gene and cDNA for human heart chymase."</b>  Urata H., Kinoshita A., Perez D.M., Misono K.S., Bumpus F.M., Graham R.M., Husain A.  J. Biol. Chem. 266:17173-17179(1991) [PubMed: 1894611] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA / MRNA].  <u>Tissue:</u> Heart.</p>
[3]	<p><b>"Determination of the primary structures of human skin chymase and cathepsin G from cutaneous mast cells of urticaria pigmentosa lesions."</b>  Schechter N.M., Wang Z.M., Blacher R.W., Lessin S.R., Lazarus G.S., Rubin H.  J. Immunol. 152:4062-4069(1994) [PubMed: 8144971] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA], PROTEIN SEQUENCE OF 22-56; 123-132; 136-148; 167-194 AND 197-247, TISSUE SPECIFICITY.  <u>Tissue:</u> Skin.</p>
[4]	<p><b>"Human protein factory for converting the transcriptome into an in vitro-expressed proteome."</b>  Goshima N., Kawamura Y., Fukumoto A., Miura A., Honma R., Satoh R., Wakamatsu A., Yamamoto J., Kimura K., Nishikawa T., Andoh T., Iida Y., Ishikawa K., Ito E., Kagawa N., Kaminaga C., Kanehori K., Kawakami B.  Nomura N.  Nat. Methods 5:1011-1017(2008) [PubMed: 19054851] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].</p>
[5]	<p>Mural R.J., Istrail S., Sutton G.G., Florea L., Halpern A.L., Mobarry C.M., Lippert R., Walenz B., Shatkay H., Dew I., Miller J.R., Flanigan M.J., Edwards N.J., Bolanos R., Fasulo D., Halldorsson B.V., Hannenhalli S., Turner R.  Venter J.C.  Submitted (SEP-2005) to the EMBL/GenBank/DDBJ databases  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].</p>
[6]	<p><b>"The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC)."</b>  The MGC Project Team  Genome Res. 14:2121-2127(2004) [PubMed: 15489334] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], VARIANT ARG-66.</p>
[7]	<p><b>"Purification and molecular cloning of chymase from human tonsils."</b>  Sukenaga Y., Kido H., Neki A., Enomoto M., Ishida K., Takagi K., Katunuma N.  FEBS Lett. 323:119-122(1993) [PubMed: 8495723] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] OF 22-247.</p>
[8]	<p><b>"Angiotensin II-forming heart chymase is a mast-cell-specific enzyme."</b>  Jenne D.E., Tschopp J.  Biochem. J. 276:567-568(1991) [PubMed: 2049082] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 26-60.  <u>Tissue:</u> Placenta.</p>
[9]	<p><b>"Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry."</b></p>

	Chen R., Jiang X., Sun D., Han G., Wang F., Ye M., Wang L., Zou H. J. Proteome Res. 8:651-661(2009) [PubMed: 19159218] [Abstract] <u>Cited for:</u> GLYCOSYLATION [LARGE SCALE ANALYSIS] AT ASN-80 AND ASN-103, MASS SPECTROMETRY. <u>Tissue:</u> Liver.
[10]	<b>"Crystal structure of phenylmethanesulfonyl fluoride-treated human chymase at 1.9 Å."</b> McGrath M.E., Mirzadegan T., Schmidt B.F. Biochemistry 36:14318-14324(1997) [PubMed: 9400368] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (1.9 ÅNGSTROMS).
[11]	<b>"The 2.2-Å crystal structure of human chymase in complex with succinyl-Ala-Ala-Pro-Phe-chloromethylketone: structural explanation for its dipeptidyl carboxypeptidase specificity."</b> Pereira P.J.P., Wang Z.-M., Rubin H., Huber R., Bode W., Schechter N.M., Strobl S. J. Mol. Biol. 286:163-173(1999) [PubMed: 9931257] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (2.2 ÅNGSTROMS).
[12]	<b>Erratum</b> Pereira P.J.P., Wang Z.-M., Rubin H., Huber R., Bode W., Schechter N.M., Strobl S. J. Mol. Biol. 286:817-817(1999) [PubMed: 10208809] [Abstract]
+	Additional computationally mapped references.

## Cross-references

### Sequence databases

- ⊙ EMBL
- ⊙ GenBank
- ⊙ DDBJ

M64269 Genomic DNA. Translation: AAA52020.1.  
M69137 Genomic DNA. Translation: AAA52021.1.  
M69136 mRNA. Translation: AAA52019.1.  
AB451464 mRNA. Translation: BAG70278.1.  
CH471078 Genomic DNA. Translation: EAW66007.1.  
BC069110 mRNA. Translation: AAH69110.1.  
BC069370 mRNA. Translation: AAH69370.1.  
BC069490 mRNA. Translation: AAH69490.1.  
BC103975 mRNA. Translation: AAI03976.1.  
S61334 mRNA. Translation: AAB26828.1.  
X59072 Genomic DNA. Translation: CAA41796.1.

IPI	IPI00013937.
PIR	KYHUCM. A40967.
RefSeq	NP_001827.1. NM_001836.2.
UniGene	Hs.135626.

### 3D structure databases

- ⊙ PDBe
- ⊙ RCSB PDB
- ⊙ PDBj

Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
1KLT	X-ray	1.90	A	22-247	[»]
1NN6	X-ray	1.75	A	20-247	[»]
1PJP	X-ray	2.20	A	22-247	[»]
1T31	X-ray	1.90	A	22-247	[»]
2HVX	X-ray	2.60	A	22-247	[»]
3N7O	X-ray	1.80	A	22-247	[»]

ProteinModelPortal	P23946.
SMR	P23946. Positions 24-247.
ModBase	Search...

**Protein-protein interaction databases**

STRING	P23946.
--------	---------

**Protein family/group databases**

MEROPS	S01.140.
--------	----------

**PTM databases**

PhosphoSite	P23946.
-------------	---------

**Proteomic databases**

PRIDE	P23946.
-------	---------

**Genome annotation databases**

Ensembl	ENST00000206446; ENSP00000206446; ENSG00000092009. ENST00000250378; ENSP00000250378; ENSG00000092009.
---------	--

GeneID	1215.
--------	-------

KEGG	hsa:1215.
------	-----------

UCSC	uc001wpp.1. human.
------	--------------------

**Organism-specific databases**

CTD	1215.
-----	-------

GeneCards	GC14M005089.
-----------	--------------

H-InvDB	HIX0202104.
---------	-------------

HGNC	HGNC:2097. CMA1.
------	------------------

HPA	CAB000363.
-----	------------

MIM	118938. gene.
-----	---------------

neXtProt	NX_P23946.
----------	------------

PharmGKB	PA26623.
----------	----------

GenAtlas	Search...
----------	-----------

**Phylogenomic databases**

eggNOG	prNOG07867.
--------	-------------

GeneTree	ENSGT00580000081281.
----------	----------------------

HOGENOM	HBG755338.
---------	------------

HOVERGEN	HBG013304.
----------	------------

InParanoid	P23946.
------------	---------

OMA	HHDIMLL.
-----	----------

OrthoDB	EOG4R23VK.
---------	------------

PhylomeDB	P23946.
-----------	---------

**Enzyme and pathway databases**

BRENDA	3.4.21.39. 247.
--------	-----------------



**Gene expression databases**

ArrayExpress	P23946.
Bgee	P23946.
CleanEx	HS_CMA1.
Genevestigator	P23946.
GermOnline	ENSG00000092009. Homo sapiens.


**Family and domain databases**

InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...

**Other Resources**

BindingDB	P23946.
NextBio	5003.
PMAP-CutDB	P23946.
SOURCE	Search...

**Entry information**

Entry name	CMA1_HUMAN
Accession	Primary (citable) accession number: <b>P23946</b> Secondary accession number(s): B5BUM8  Q9UDH5
Entry history	Integrated into March 1, 1992 UniProtKB/Swiss-Prot: Last sequence update: March 1, 1992 Last modified: April 5, 2011 This is version 118 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
Disclaimer	Any medical or genetic information present in this entry is provided for research, educational and informational purposes

only. It is not in any way intended to be used as a substitute for professional medical advice, diagnosis, treatment or care.

### Relevant documents

Human chromosome 14

Human chromosome 14: entries, gene names and cross-references to MIM

Human entries with polymorphisms or disease mutations

List of human entries with polymorphisms or disease mutations

Human polymorphisms and disease mutations

Index of human polymorphisms and disease mutations

MIM cross-references

Online Mendelian Inheritance in Man (MIM) cross-references in UniProtKB/Swiss-Prot

PDB cross-references

Index of Protein Data Bank (PDB) cross-references

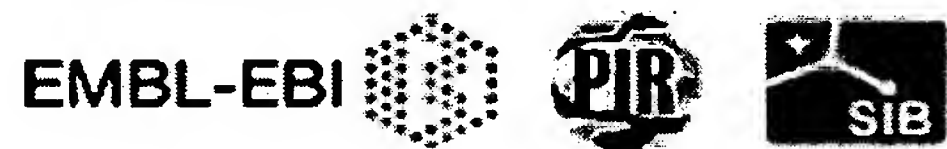
Peptidase families

Classification of peptidase families and list of entries

SIMILARITY comments

Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



Protein

Translations of Life

Display Settings: GenPept

chymase [Homo sapiens]

GenBank: AAB26828.1

[FASTA](#) [Graphics](#)

[Go to:](#)

```
LOCUS      AAB26828          226 aa          linear    PRI 19-OCT-1993
DEFINITION chymase [Homo sapiens].
ACCESSION  AAB26828
VERSION    AAB26828.1  GI:409011
DBSOURCE   accession S61334.1
KEYWORDS   .
SOURCE     Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE  1 (residues 1 to 226)
AUTHORS    Sukenaga,Y., Kido,H., Neki,A., Enomoto,M., Ishida,K., Takagi,K. and
            Katunuma,N.
TITLE      Purification and molecular cloning of chymase from human tonsils
JOURNAL    FEBS Lett. 323 (1-2), 119-122 (1993)
PUBMED     8495723
REMARK     GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 132634] from the original journal article.
COMMENT    Method: conceptual translation supplied by author.
FEATURES   Location/Qualifiers
     source          1..226
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     Protein         1..226
                     /product="chymase"
     Region          1..222
                     /region_name="Tryp_SpC"
                     /note="Trypsin-like serine protease; Many of these are
                     synthesized as inactive precursor zymogens that are
                     cleaved during limited proteolysis to generate their
                     active forms. Alignment contains also inactive enzymes
                     that have substitutions of the catalytic...; cd00190"
                     /db_xref="CDD:29152"
     Site            1
                     /site_type="cleavage"
                     /db_xref="CDD:29152"
     Site            order(45,89,182)
                     /site_type="active"
                     /db_xref="CDD:29152"
     Site            order(176,197,199)
                     /site_type="other"
                     /note="substrate binding sites"
                     /db_xref="CDD:29152"
     CDS            1..226
                     /gene="chymase"
                     /coded_by="S61334.1:1..681"
                     /note="Sequence start is N-terminal of purified protein"
ORIGIN
  1 iiggteskph srpymaylei vtsngpskfc ggflirrnfv ltaahcagrs itvtlgahni
  61 teeedtwqkl evikqfrhpk yntstlhhdh mllklkekas ltlavgtlpf psqfnfvppg
 121 rmcrvagwgr tgvlkpgsd t lgevkrlrmd pqacshfrdf dhnqlqlcvgn prktksafkg
 181 dsggpilcag vaqgivsygr sdakppavft rishyrpwin qilqan
//
```

## Protein

Translations of Life

Display Settings: GenPept

## mast cell chymase [Homo sapiens]

GenBank: AAA52020.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAA52020 247 aa linear PRI 01-NOV-1994  
DEFINITION mast cell chymase [Homo sapiens].  
ACCESSION AAA52020  
VERSION AAA52020.1 GI:180542  
DBSOURCE locus HUMCHYMASE accession M64269.1  
KEYWORDS .  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 247)  
AUTHORS Caughey,G.H., Zerweck,E.H. and Vanderslice,P.  
TITLE Structure, chromosomal assignment, and deduced amino acid sequence  
of a human gene for mast cell chymase  
J. Biol. Chem. 266 (20), 12956-12963 (1991)  
PUBMED 2071582  
COMMENT Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..247  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="Unassigned"  
/cell\_type="mast cell"  
/tissue\_type="placenta"  
Protein 1..247  
/product="mast cell chymase"  
sig\_peptide 1..21  
/gene="CMA1"  
/note="G00-127-603"  
mat\_peptide 22..247  
/gene="CMA1"  
/product="mast cell chymase"  
/note="G00-127-603"  
Region 22..243  
/region\_name="Tryp\_Spc"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cd00190"  
/db\_xref="CDD:29152"  
Site 22  
/site\_type="cleavage"  
/db\_xref="CDD:29152"  
Site order(66,110,203)  
/site\_type="active"  
/db\_xref="CDD:29152"  
Site order(197,218,220)  
/site\_type="other"  
/note="substrate binding sites"  
/db\_xref="CDD:29152"  
CDS 1..247  
/gene="CMA1"  
/coded\_by="join(M64269.1:5158..5215,M64269.1:5886..6036,  
M64269.1:6786..6921,M64269.1:7108..7362,  
M64269.1:7733..7876)"  
/db\_xref="GDB:G00-127-603"  
ORIGIN  
1 mlllppllll flcsraeag eiiggteckp hsrpymayle ivtsngpskf cggflirrf  
61 vltaahcagr sitvtlgahn iteedtwqk levikqfrhp kyntstlhhd imllklkeka  
121 sltlavgtlp fpsqfnfvpp grmcrvagwg rtgvlkpgsd tlgevklrlm dpqacshfrd  
181 fdhnlqlcvg nprktsafk gdsggplla gvaqgivsyg rsdakppavf trishyrpwi  
241 nqilqan  
//



## Protein

Translations of Life

Display Settings: GenPept

## chymase [Homo sapiens]

GenBank: AAA52021.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAA52021 247 aa linear PRI 01-NOV-1994  
DEFINITION chymase [Homo sapiens].  
ACCESSION AAA52021  
VERSION AAA52021.1 GI:180544  
DBSOURCE locus HUMCHYMB accession M69137.1  
KEYWORDS .  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 247)  
AUTHORS Urata,H., Kinoshita,A., Perez,D.M., Misono,K.S., Bumpus,F.M.,  
Graham,R.M. and Husain,A.  
TITLE Cloning of the gene and cDNA for human heart chymase  
JOURNAL J. Biol. Chem. 266 (26), 17173-17179 (1991)  
PUBMED 1894611  
COMMENT Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..247  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="Unassigned"  
/tissue\_type="heart # 117"  
Protein 1..247  
/product="chymase"  
/EC\_number="3.4.21.39"  
Region 22..243  
/region\_name="Tryp\_SpC"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cd00190"  
/db\_xref="CDD:29152"  
Site 22  
/site\_type="cleavage"  
/db\_xref="CDD:29152"  
Site order(66,110,203)  
/site\_type="active"  
/db\_xref="CDD:29152"  
Site order(197,218,220)  
/site\_type="other"  
/note="substrate binding sites"  
/db\_xref="CDD:29152"  
CDS 1..247  
/gene="CMA1"  
/standard\_name="angiotensin I convertase"  
/coded\_by="join(M69137.1:259..316,M69137.1:989..1139,  
M69137.1:1886..2021,M69137.1:2208..2462,  
M69137.1:2835..2978)"  
/db\_xref="GDB:G00-127-603"  
ORIGIN  
1 mlllpplp111 fl1csraeag eiiggteckp hsrpymayle ivtsngpskf cggflirnf  
61 vltaaacagr sitvtlgahn iteedtwqk levikqfrhp kyntstlhhd imllklkeka  
121 sltlavgtlp fpsqfnfvpp grmcrvagwg rtgvlpkgsd tlqevklrlm dpqacshfrd  
181 fdhnlqlcvg nprktsafk gdsqgp1lca gvaqgivsyt rsdakppavf trishyprwi  
241 nqilqan  
//

## Protein

Translations of Life

Display Settings: GenPept

## chymase [Homo sapiens]

GenBank: AAA52019.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAA52019 247 aa linear PRI 01-NOV-1994  
DEFINITION chymase [Homo sapiens].  
ACCESSION AAA52019  
VERSION AAA52019.1 GI:180540  
DBSOURCE locus HUMCHYMA accession M69136.1  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 247)  
AUTHORS Urata,H., Kinoshita,A., Perez,D.M., Misono,K.S., Bumpus,F.M.,  
Graham,R.M. and Husain,A.  
TITLE Cloning of the gene and cDNA for human heart chymase  
JOURNAL J. Biol. Chem. 266 (26), 17173-17179 (1991)  
PUBMED 1894611  
COMMENT Method: conceptual translation.  
FEATURES  
source 1..247  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="Unassigned"  
/tissue\_type="heart # 117"  
Protein 1..247  
/product="chymase"  
/EC\_number="3.4.21.39"  
Region 22..243  
/region\_name="Tryp\_SpC"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cd00190"  
/db\_xref="CDD:29152"  
Site 22  
/site\_type="cleavage"  
/db\_xref="CDD:29152"  
Site order(66,110,203)  
/site\_type="active"  
/db\_xref="CDD:29152"  
Site order(197,218,220)  
/site\_type="other"  
/note="substrate binding sites"  
/db\_xref="CDD:29152"  
CDS 1..247  
/gene="CMA1"  
/standard\_name="angiotensin I convertase"  
/coded\_by="M69136.1:16..759"  
/db\_xref="GDB:G00-127-603"

ORIGIN  
1 mlllplplll flfcsraeag eiiggteckp hsrpymayle ivtsngpskf cggflirrf  
61 vltaaacagr sitvtlgahn iteedtwgk levikqfrhp kyntstlhhd imllklkeka  
121 sltlavgtlp fspqfnfvpp grmcrvagwg rtgvlpkgsd tlqevklrlm dpqacshfrd  
181 fdhnlqlcvg nprktsafk gdsggpllca gvaqgivsyg rsdakppavf trishyprwi  
241 nqilqan  
//

**P21842 (CMA1\_CANFA) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified March 8, 2011. Version 79.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Chymase</b> EC=3.4.21.39 <i>Alternative name(s):</i> Alpha-chymase Mast cell protease I
Gene names	Name: <b>CMA1</b>
Organism	<b>Canis familiaris (Dog) (Canis lupus familiaris)</b>
Taxonomic identifier	9615 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Laurasiatheria › Carnivora › Caniformia › Canidae › Canis

**Protein attributes**

Sequence length	249 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**











Function	Major secreted protease of mast cells with suspected roles in vasoactive peptide generation, extracellular matrix degradation, and regulation of gland secretion.
Catalytic activity	Preferential cleavage: Phe- -Xaa > Tyr- -Xaa > Trp- -Xaa > Leu- -Xaa.
Subcellular location	Secreted <a href="#">By similarity</a> . Cytoplasmic granule <a href="#">By similarity</a> . Note: Mast cell granules <a href="#">By similarity</a> .
Sequence similarities	Belongs to the peptidase S1 family. Granzyme subfamily. Contains 1 peptidase S1 domain.

**Ontologies****Keywords**

Cellular component	Secreted
Domain	Signal

Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Zymogen
Technical term	Direct protein sequencing
Gene Ontology (GO)	
Biological process	proteolysis Inferred from electronic annotation. Source: InterPro
Complete GO annotation...	

Sequence annotation (Features)


Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
Molecule processing					
	Signal peptide	1 – 19	19		
	Propeptide	20 – 21	2	Activation peptide	PRO_000002743
	Chain	22 – 249	228	Chymase	PRO_000002743
Regions					
	Domain	22 – 245	224	Peptidase S1	
Sites					
	Active site	66	1	Charge relay system By similarity	
	Active site	110	1	Charge relay system By similarity	
	Active site	203	1	Charge relay system By similarity	
Amino acid modifications					
	Disulfide bond	51 ↔ 67		By similarity	
	Disulfide bond	144 ↔ 209		By similarity	
	Disulfide bond	175 ↔ 188		By similarity	



Experimental info

<input type="checkbox"/>	Sequence conflict	29	1	K → R AA sequence Ref.3		
<input type="checkbox"/>	Sequence conflict	38	1	H → Q AA sequence Ref.3		
<input type="checkbox"/>	Sequence conflict	45	1	R → T AA sequence Ref.3		

Sequences

Sequence	Length	Mass (Da)
 P21842 [UniParc]. Last modified May 1, 1991. Version 1. Checksum: 3BBD0A6C2855F540	FASTA	249 27,812
<div><div>102030405060 MHCLPLTLL LLLCSRAEAE EIIGGTESKP HSRPYMAHLE ILTLRNHLAS CGGFLIRRNF 708090100110120 VLTAHCAGR FIMVTLGAHN IQKKEDTWQK LEVIKQFPHP KYDDLTLRHD IMLLKLKEKA 130140150160170180 NLTLAVGTLP LSPQFNFVPP GRMCRVAGWG KRQVNGSGSD TLQEVKLRLM DPQACRHYMA 190200210220230240 FDHNLQLCVG NPRKTKSAFK GDSGGPLLCA GVAQGIVSYG QNDAKPPAVF TRISHYRPWI NKVLKQNKAK</div><div>« Hide</div></div>		

References

[1]	<p><b>"Dog mast cell chymase: molecular cloning and characterization."</b> Caughey G.H., Raymond W.W., Vanderslice P. Biochemistry 29:5166-5171(1990) [PubMed: 2378872] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].</p>
[2]	<p><b>"Cloning and expression of the dog mast cell alpha-chymase gene."</b> Caughey G.H., Blount J.L., Koerber K.L., Kitamura M., Fang K.C. J. Immunol. 159:4367-4375(1997) [PubMed: 9379034] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA]. <u>Strain:</u> Beagle.</p>
[3]	<p><b>"Purification and characterization of dog mastocytoma chymase: identification of an octapeptide conserved in chymotryptic leukocyte proteinases."</b> Caughey G.H., Viro N.F., Lazarus S.C., Nadel J.A. Biochim. Biophys. Acta 952:142-149(1988) [PubMed: 3122835] [Abstract] <u>Cited for:</u> PROTEIN SEQUENCE OF 22-46.</p>

**Cross-references****Sequence databases**

Ⓢ EMBL	J02904 mRNA. Translation: AAA30835.1.
Ⓢ GenBank	U89607 Genomic DNA. Translation: AAB94641.1.
Ⓢ DDBJ	

PIR	A35842.
-----	---------

RefSeq	NP_001013442.1. NM_001013424.1. XP_855610.1. XM_850517.1.
--------	--

UniGene	Cfa.16336.
---------	------------

**3D structure databases**

ProteinModelPortal	P21842.
--------------------	---------

SMR	P21842. Positions 22-247.
-----	---------------------------

ModBase	Search...
---------	-----------

**Protein-protein interaction databases**

STRING	P21842.
--------	---------

**Protein family/group databases**

MEROPS	S01.140.
--------	----------

**Genome annotation databases**

Ensembl	ENSCAFT00000019746; ENSCAFP00000018316; ENSCAFG00000012443.
---------	--

GeneID	490628.
--------	---------

KEGG	cfa:490628.
------	-------------

**Organism-specific databases**

CTD	490628.
-----	---------

**Phylogenomic databases**

eggNOG	maNOG16150.
--------	-------------

GeneTree	ENSGT00580000081281.
----------	----------------------

HOVERGEN	HBG013304.
----------	------------

InParanoid	P21842.
------------	---------

OrthoDB	EOG4R23VK.
---------	------------

PhylomeDB	P21842.
-----------	---------

**Enzyme and pathway databases**

BRENDA	3.4.21.39. 463.
--------	-----------------

**Family and domain databases**

InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...

**Entry information**

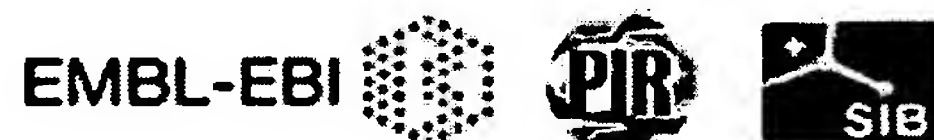
Entry name	CMA1_CANFA
Accession	Primary (citable) accession number: <b>P21842</b>
Entry history	Integrated into May 1, 1991 UniProtKB/Swiss-Prot: Last sequence May 1, 1991 update: Last modified: March 8, 2011 This is version 79 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program

**Relevant documents**

Peptidase families  
Classification of peptidase families and list of entries

SIMILARITY comments  
Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



**P21844 (CMA1\_MOUSE) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified March 8, 2011. Version 102.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Chymase</b> EC=3.4.21.39 <i>Alternative name(s):</i> Alpha-chymase Mast cell chymase 1 Mast cell protease 5 Short name=mMCP-5 Mast cell protease I
Gene names	Name: <b>Cma1</b> Synonyms: Mcpt5
Organism	<b>Mus musculus (Mouse)</b>
Taxonomic identifier	10090 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires › Rodentia › Sciurognathi › Muroidea › Muridae › Murinae › Mus › Mus

**Protein attributes**

Sequence length	247 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Major secreted protease of mast cells with suspected roles in vasoactive peptide generation, extracellular matrix degradation, and regulation of gland secretion.
Catalytic activity	Preferential cleavage: Phe- -Xaa > Tyr- -Xaa > Trp- -Xaa > Leu- -Xaa.
Subcellular location	Secreted. Cytoplasmic granule. Note: Secretory granules. <span style="border: 1px solid black; padding: 0 2px;">Ref.6</span>
Tissue specificity	Mast cells.
Sequence similarities	Belongs to the peptidase S1 family. Granzyme subfamily. Contains 1 peptidase S1 domain.
Sequence caution	The sequence CAA48705.1 differs from that shown. Reason: Erroneous initiation.



Ontologies

Keywords

Cellular component	Secreted
Domain	Signal
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Glycoprotein Zymogen
Technical term	Direct protein sequencing

Gene Ontology (GO)

Biological process	proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	intracellular Inferred from direct assay. Source: MGI

Complete GO annotation...

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
-------------	-------------	--------	-------------	----------------	--------------------

Molecule processing

<input checked="" type="checkbox"/>	Signal peptide	1 – 19	19	Potential		
<input type="checkbox"/>	Propeptide	20 – 21	2	Activation peptide		PRO_000002745
<input checked="" type="checkbox"/>	Chain	22 – 247	226	Chymase		PRO_000002745

Regions


<input checked="" type="checkbox"/>	Domain	22 – 245	224	Peptidase S1		
-------------------------------------	--------	----------	-----	--------------	--	--

Sites

<input type="checkbox"/>	Active site	66	1	Charge relay system By similarity		
<input type="checkbox"/>	Active site	110	1	Charge relay system By similarity		

<input type="checkbox"/>	Active site	203	1	Charge relay system <div>By similarity</div>		
<b>Amino acid modifications</b>						
<input type="checkbox"/>	Glycosylation	80	1	N-linked (GlcNAc...) <div>Potential</div>		
<input type="checkbox"/>	Disulfide bond	51 ↔ 67		<div>By similarity</div>		
<input type="checkbox"/>	Disulfide bond	144 ↔ 209		<div>By similarity</div>		
<input type="checkbox"/>	Disulfide bond	175 ↔ 188		<div>By similarity</div>		
<b>Experimental info</b>						
<input type="checkbox"/>	Sequence conflict	5	1	T → A in AAA40105. <div>Ref.1</div>		
<input type="checkbox"/>	Sequence conflict	51	1	C → R AA sequence <div>Ref.5</div>		
<input type="checkbox"/>	Sequence conflict	224	1	A → R in AAA39492. <div>Ref.3</div>		

## Sequences

**Sequence** **Length** **Mass (Da)**  
 **P21844 [UniParc].** **FASTA** **247** **27,586**  
 Last modified December 1, 1992. Version 2.  
 Checksum: 24C290CF61237DC7

```

      10      20      30      40      50      60
MHLLTLHLLL LLLGSSTKAG EIIGGTECIP HSRPYMAYLE IVTSENYLSA CSGFLIRRNF

      70      80      90     100     110     120
VLTAAHCAGR SITVLLGAHN KTSKEDTWQK LEVEKQFLHP KYDENLVVHD IMLLKLKEKA

     130     140     150     160     170     180
KLTLGVGTLP LSAFNFIIPP GRMCRAVGWG RTNVNEPASD TLQEVKMRLQ EPQACKHFTS

     190     200     210     220     230     240
FRHNSQLCVG NPKKMQNVYK GDSGGPLLCA GIAQGLASYV HRNAKPPAVF TRISHYRPWI

NKILREN
  
```

« Hide

## References

- [1] "Molecular cloning of the mouse mast cell protease-5 gene. A novel secretory granule protease expressed early in the differentiation of serosal mast cells."

	McNeil H.P., Austen K.F., Somerville L.L., Gurish M.F., Stevens R.L. J. Biol. Chem. 266:20316-20322(1991) [PubMed: 1939089] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].
[2]	<b>"Cloning and structural analysis of MMCP-1, MMCP-4 and MMCP-5, three mouse mast cell-specific serine proteases."</b> Huang R., Blom T., Hellman L. Eur. J. Immunol. 21:1611-1621(1991) [PubMed: 2060576] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA]. <u>Strain:</u> Leaden X A1. <u>Tissue:</u> Mastocytoma.
[3]	<b>"Molecular cloning and characterization of mouse mast cell chymases."</b> Chu W., Johnson D.A., Musich P.R. Biochim. Biophys. Acta 1121:83-87(1992) [PubMed: 1376147] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].
[4]	<b>"Characterization of the gene encoding mouse mast cell protease 8 (mMCP-8), and a comparative analysis of hematopoietic serine protease genes."</b> Lunderius C., Hellman L. Immunogenetics 53:225-232(2001) [PubMed: 11398967] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-19. <u>Strain:</u> 129/Sv.
[5]	<b>"Different mouse mast cell populations express various combinations of at least six distinct mast cell serine proteases."</b> Reynolds D.S., Stevens R.L., Lane W.S., Carr M.H., Austen K.F., Serafin W.E. Proc. Natl. Acad. Sci. U.S.A. 87:3230-3234(1990) [PubMed: 2326280] [Abstract] <u>Cited for:</u> PROTEIN SEQUENCE OF 22-51.
[6]	<b>"Translation and granule localization of mouse mast cell protease-5. Immunodetection with specific antipeptide Ig."</b> McNeil H.P., Frenkel D.P., Austen F., Friend D.S., Stevens R.L. J. Immunol. 149:2466-2472(1992) [PubMed: 1527387] [Abstract] <u>Cited for:</u> SUBCELLULAR LOCATION.
+	Additional computationally mapped references.

## Cross-references

### Sequence databases

⊙ EMBL	M73759 mRNA. Translation: AAA40105.1.
⊙ GenBank	M73760 mRNA. No translation available.
⊙ DDBJ	X68805 mRNA. Translation: CAA48705.1. Different initiation.
	M68898 mRNA. Translation: AAA39492.1.
	AF119364 Genomic DNA. Translation: AAD43901.1.

IPI	IPI00230426.
PIR	S23504. S26043.
RefSeq	NP_034910.1. NM_010780.2.
UniGene	Mm.1252.

### 3D structure databases

ProteinModelPortal	P21844.
SMR	P21844. Positions 22-247.

ModBase	Search...
<b>Protein-protein interaction databases</b>	
STRING	P21844.
<b>Protein family/group databases</b>	
MEROPS	S01.150.
<b>PTM databases</b>	
PhosphoSite	P21844.
<b>Genome annotation databases</b>	
Ensembl	ENSMUST00000022834; ENSMUSP00000022834; ENSMUSG00000022225.
GeneID	17228.
KEGG	mmu:17228.
UCSC	uc007ubf.1. mouse.
<b>Organism-specific databases</b>	
CTD	17228.
MGI	MGI:96941. Cma1.
<b>Phylogenomic databases</b>	
GeneTree	ENSGT00580000081281.
HOVERGEN	HBG013304.
InParanoid	P21844.
OrthoDB	EOG4R23VK.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.21.39. 244.
<b>Gene expression databases</b>	
ArrayExpress	P21844.
Bgee	P21844.
CleanEx	MM_CMA1.
Genevestigator	P21844.
GermOnline	ENSMUSG00000022225. Mus musculus.
<b>Family and domain databases</b>	
InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]



Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
NextBio	291648.
SOURCE	Search...
<b>Entry information</b>	
Entry name	CMA1_MOUSE
Accession	Primary (citable) accession number: <b>P21844</b> Secondary accession number(s): Q9R1F0
Entry history	Integrated into May 1, 1991 UniProtKB/Swiss-Prot: Last sequence December 1, 1992 update: Last modified: March 8, 2011 This is version 102 of the entry and version 2 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Relevant documents</b>	
MGD cross-references Mouse Genome Database (MGD) cross-references in UniProtKB/Swiss-Prot	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License &amp; Disclaimer | Contact



**P56435 (CMA1\_MACFA) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified January 11, 2011. Version 60.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Chymase</b> EC=3.4.21.39 <i>Alternative name(s):</i> Alpha-chymase
Gene names	Name: <b>CMA1</b>
Organism	<b>Macaca fascicularis (Crab-eating macaque) (Cynomolgus monkey)</b>
Taxonomic identifier	9541 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Cercopithecidae › Cercopithecinae › Macaca

**Protein attributes**

Sequence length	247 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at transcript level.

**General annotation (Comments)**








Function	Major secreted protease of mast cells with suspected roles in vasoactive peptide generation, extracellular matrix degradation, and regulation of gland secretion <a href="#">By similarity</a> .
Catalytic activity	Preferential cleavage: Phe- -Xaa > Tyr- -Xaa > Trp- -Xaa > Leu- -Xaa.
Subcellular location	Secreted <a href="#">By similarity</a> . Cytoplasmic granule <a href="#">By similarity</a> . Note: Mast cell granules <a href="#">By similarity</a> .
Sequence similarities	Belongs to the peptidase S1 family. Granzyme subfamily. Contains 1 peptidase S1 domain.

**Ontologies**

<b>Keywords</b>	
Cellular component	Secreted
Domain	Signal

Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Glycoprotein Zymogen
<b>Gene Ontology (GO)</b>	
Biological process	proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	extracellular region Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro
Complete GO annotation...	

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
<b>Molecule processing</b>					
 Signal peptide	1 – 19	19	By similarity		
 Propeptide	20 – 21	2	Activation peptide		PRO_0000027435
 Chain	22 – 247	226	Chymase		PRO_0000027436
<b>Regions</b>					
 Domain	22 – 245	224	Peptidase S1		
<b>Sites</b>					
 Active site	66	1	Charge relay system By similarity		
 Active site	110	1	Charge relay system By similarity		
 Active site	203	1	Charge relay system By similarity		

Amino acid modifications

<input type="checkbox"/>	Glycosylation	80	1	N-linked (GlcNAc...) Potential		
<input type="checkbox"/>	Glycosylation	103	1	N-linked (GlcNAc...) Potential		
<input type="checkbox"/>	Disulfide bond	51 ↔ 67		By similarity		
<input type="checkbox"/>	Disulfide bond	144 ↔ 209		By similarity		
<input type="checkbox"/>	Disulfide bond	175 ↔ 188		By similarity		

Sequences

Sequence	Length	Mass (Da)
<input checked="" type="checkbox"/> P56435 [UniParc]. Last modified July 15, 1998. Version 1. Checksum: 6D049B34377FC8B9	FASTA 247	27,400
<div><div><div><div>102030405060</div><div>MLLLPLPLL FFLCSRAEAG EIIGGTECKP HSRPYMAYLE IVTSNGPSKS CGGFLIRRNF</div></div><div><div>708090100110120</div><div>VLTAVHCAGR SITVTLAGHN ITEKEDTWQK LEVIKQFRHP KYNTSTLHHD IMLLKLKEKA</div></div><div><div>130140150160170180</div><div>SLTLAVGTLP FPSQFNFVPP GRMCRVAGWG RTGVLKPGSD TLQEVKLRLM DPQACSHFRY</div></div><div><div>190200210220230240</div><div>FDHNLQLCVG NPRKTKSAFK GDSGGPLLCA GVAQGIVSYG RLDAKPPAVF TRISHYRPWI</div></div><div>NKILQAN</div></div><div>« Hide</div></div>		

References

[1] "Induction of chymase that forms angiotensin II in the monkey atherosclerotic aorta."  
Takai S., Shiota N., Kobayashi S., Matsumura E., Miyazaki M.  
FEBS Lett. 412:86-90(1997) [PubMed: 9257695] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA].

Cross-references

Sequence databases

- EMBL
- GenBank
- DDBJ

AB000823 mRNA. Translation: BAA22070.1.



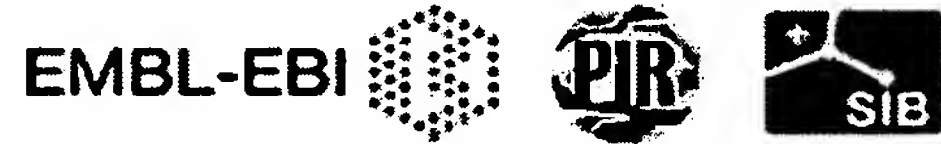
<b>3D structure databases</b>	
ProteinModelPortal	P56435.
SMR	P56435. Positions 22-247.
ModBase	Search...
<b>Protein family/group databases</b>	
MEROPS	S01.140.
<b>Phylogenomic databases</b>	
HOVERGEN	HBG013304.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.21.39. 3438.
<b>Family and domain databases</b>	
InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. False negative. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Entry information</b>	
Entry name	CMA1_MACFA
Accession	Primary (citable) accession number: <b>P56435</b>
Entry history	Integrated into July 15, 1998 UniProtKB/Swiss-Prot: Last sequence update: July 15, 1998 Last modified: January 11, 2011 This is version 60 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program

**Relevant documents**

Peptidase families  
Classification of peptidase families and list of entries

SIMILARITY comments  
Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



**P52195 (CMA1\_PAPHA) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified January 11, 2011. Version 57.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Chymase</b> EC=3.4.21.39 <i>Alternative name(s):</i> Alpha-chymase Mast cell chymase
Gene names	Name: <b>CMA1</b> Synonyms: CHM
Organism	<b>Papio hamadryas (Hamadryas baboon)</b>
Taxonomic identifier	9557 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Cercopithecidae › Cercopithecinae › Papio

**Protein attributes**

Sequence length	247 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at transcript level.

**General annotation (Comments)**












Function	Major secreted protease of mast cells with suspected roles in vasoactive peptide generation, extracellular matrix degradation, and regulation of gland secretion.
Catalytic activity	Preferential cleavage: Phe- -Xaa > Tyr- -Xaa > Trp- -Xaa > Leu- -Xaa.
Subcellular location	Secreted. Cytoplasmic granule. Note: Mast cell granules.
Sequence similarities	Belongs to the peptidase S1 family. Granzyme subfamily. Contains 1 peptidase S1 domain.

**Ontologies**

<b>Keywords</b>	
Cellular component	Secreted
Domain	Signal

Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Glycoprotein Zymogen
<b>Gene Ontology (GO)</b>	
Biological process	proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	extracellular region Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro
Complete GO annotation...	

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
<b>Molecule processing</b>					
	Signal peptide	1 – 19	19		
	Propeptide	20 – 21	2	Activation peptide	PRO_0000027437
	Chain	22 – 247	226	Chymase	PRO_0000027438
<b>Regions</b>					
	Domain	22 – 245	224	Peptidase-S1	
<b>Sites</b>					
	Active site	66	1	Charge relay system 	
	Active site	110	1	Charge relay system 	
	Active site	203	1	Charge relay system 	



**Amino acid modifications**

<input type="checkbox"/>	Glycosylation	80	1	N-linked (GlcNAc...) Potential		
<input type="checkbox"/>	Glycosylation	103	1	N-linked (GlcNAc...) Potential		
<input type="checkbox"/>	Disulfide bond	51 ↔ 67		By similarity		
<input type="checkbox"/>	Disulfide bond	144 ↔ 209		By similarity		
<input type="checkbox"/>	Disulfide bond	175 ↔ 188		By similarity		

**Sequences**

Sequence	Length	Mass (Da)
----------	--------	-----------

<input checked="" type="checkbox"/> P52195 [UniParc].	FASTA	247	27,339
---	-------	-----	--------

Last modified October 1, 1996. Version 1.  
Checksum: E0EC15E0FA72FD8B

```

      10      20      30      40      50      60
MLLLPLPLL LFLCSRAEAG EIIGGTECKP HSRPYMAYLE IVTSNGPSKS CGGFLIRRNF

      70      80      90     100     110     120
VLTAHCAGR SITVTLAGHN ITEKEDTWQE LEVIKQFRHP KYNTSTLHHD IMLLKLKEKA

     130     140     150     160     170     180
SLTLAVGTLF FPSQFNFPVP GRMCRVAGWG RTGVLPKPGSD TLQEVKLRLM DPQACSHFRY

     190     200     210     220     230     240
FDHNLQLCVG NPRKTKSAFK GDSGGPLLCA GVAQGIVSYG RLDAPPAVF TRISHYRPWI

NKILQAN

```

« Hide

**References**

- [1] Liao Y., Karnik S., Husain A.  
Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases  
Cited for: NUCLEOTIDE SEQUENCE [GENOMIC DNA / MRNA].

**Cross-references****Sequence databases**

<ul style="list-style-type: none"> <li>⊗ EMBL</li> <li>⊗ GenBank</li> <li>⊗ DDBJ</li> </ul>	U38521 mRNA. Translation: AAA91160.1. U38463 Genomic DNA. Translation: AAA91159.1.
---	---

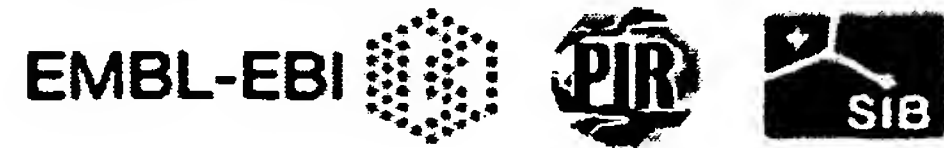
**3D structure databases**

ProteinModelPortal	P52195.
--------------------	---------

SMR	P52195. Positions 22-247.
ModBase	Search...
<b>Protein family/group databases</b>	
MEROPS	S01.140.
<b>Phylogenomic databases</b>	
HOVERGEN	HBG013304.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.21.39. 39388.
<b>Family and domain databases</b>	
InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Entry information</b>	
Entry name	CMA1_PAPHA
Accession	Primary (citable) accession number: <b>P52195</b>
Entry history	Integrated into      October 1, 1996 UniProtKB/Swiss- Prot: Last sequence      October 1, 1996 update: Last modified:      January 11, 2011 This is version 57 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Relevant documents</b>	
Peptidase families Classification of peptidase families and list of entries	

SIMILARITY comments  
Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



**P50339 (CMA1\_RAT) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified January 11, 2011. Version 75.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Chymase</b> EC=3.4.21.39 <i>Alternative name(s):</i> Alpha-chymase Mast cell protease 3 Short name=rMCP-3 Mast cell protease 5 Short name=rMCP-5 Mast cell protease III Short name=rMCP-III
Gene names	Name: <b>Cma1</b> Synonyms: Mcpt3
Organism	<b>Rattus norvegicus (Rat)</b>
Taxonomic identifier	10116 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires › Rodentia › Sciurognathi › Muroidea › Muridae › Murinae › Rattus

**Protein attributes**

Sequence length	247 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at transcript level.

**General annotation (Comments)**

Function	Major secreted protease of mast cells with suspected roles in vasoactive peptide generation, extracellular matrix degradation, and regulation of gland secretion <a href="#">By similarity</a> .
Catalytic activity	Preferential cleavage: Phe- -Xaa > Tyr- -Xaa > Trp- -Xaa > Leu- -Xaa.
Subcellular location	Secreted. Cytoplasmic granule. Note: Secretory granules.
Tissue specificity	Mast cells.
Sequence similarities	Belongs to the peptidase S1 family. Granzyme subfamily. Contains 1 peptidase S1 domain.



Ontologies

Keywords

Cellular component	Secreted
Domain	Signal
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Glycoprotein Zymogen

Gene Ontology (GO)

Biological process	proteolysis Inferred from mutant phenotype. Source: RGD
Cellular component	extracellular region Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	peptide binding Inferred from mutant phenotype. Source: RGD  serine-type endopeptidase activity Inferred from mutant phenotype. Source: RGD

Complete GO annotation...

Sequence annotation (Features)

	Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
Molecule processing						
<input checked="" type="checkbox"/>	Signal peptide	1 – 19	19	Potential		
<input type="checkbox"/>	Propeptide	20 – 21	2	Activation peptide		PRO_0000027443
<input checked="" type="checkbox"/>	Chain	22 – 247	226	Chymase		PRO_0000027444
Regions						
<input checked="" type="checkbox"/>	Domain	22 – 245	224	Peptidase S1		
Sites						
<input type="checkbox"/>	Active site	66	1	Charge relay system By similarity		
<input type="checkbox"/>	Active site	110	1	Charge relay		

<input type="checkbox"/>	Active site	203	1	<div>system<div>By similarity</div></div> <div>Charge relay system<div>By similarity</div></div>	I	
--------------------------	-------------	-----	---	--	---	--

Amino acid modifications

<input type="checkbox"/>	Glycosylation	80	1	<div>N-linked (GlcNAc...)<div>Potential</div></div>	I	
<input type="checkbox"/>	Disulfide bond	51 ↔ 67		<div>By similarity</div>	II	
<input type="checkbox"/>	Disulfide bond	144 ↔ 209		<div>By similarity</div>	I	I
<input type="checkbox"/>	Disulfide bond	175 ↔ 188		<div>By similarity</div>	II	

Sequences

Sequence	Length	Mass (Da)
<div><input checked="" type="checkbox"/> P50339 [UniParc]. Last modified October 1, 1996. Version 1. Checksum: 6525D7BF1BFDF053</div> <div><div>102030405060</div><div>MNLHALCLLL LLLGSSTKAG EIIGGTECIP HSRPYMAYLE IVTSDNYLSA CSGFLIRRNF</div><div>708090100110120</div><div>VLTAAHCAGR SITVLLGAHN KTYKEDTWQK LEVEKQFIHP NYDKRLVLHD IMLLKLKEKA</div><div>130140150160170180</div><div>KLTLGVGTLP LSAFNFIPT GRMCRAVGWG RTNVNEPASD TLQEVKMRLQ EPQSCKHFTS</div><div>190200210220230240</div><div>FQHKSQLCVG NPKKMQNVYK GDSGGPLLCA GIAQGIASYV HPNAKPPAVF TRISHYRPWI</div><div>NKILREN</div></div>	FASTA	24727,569

« Hide

References

[1]

"Cloning of the cDNA encoding a novel rat mast-cell proteinase, rMCP-3, and its expression in comparison with other rat mast-cell proteinases."  
Ide H., Itoh H., Tomita M., Murakumo Y., Kobayashi T., Maruyama H., Osada Y., Nawa Y.  
Biochem. J. 311:675-680(1995) [PubMed: 7487912] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA].  
Tissue: Peritoneal mast cell.

[2]

"Secretory granule proteases in rat mast cells. Cloning of 10 different serine proteases and a carboxypeptidase A from various rat mast cell populations."  
Lutzelschwab C., Pejler G., Aveskogh M., Hellman L.  
J. Exp. Med. 185:13-29(1997) [PubMed: 8996238] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA] OF 2-247.  
Strain: Sprague-Dawley.

+ Additional computationally mapped references.

## Cross-references

### Sequence databases

⊕ EMBL	D38495 mRNA. Translation: BAA07507.1.
⊕ GenBank	U67908 mRNA. Translation: AAB48261.1.
⊕ DDBJ	
IPI	IPI00197397.
PIR	S59135.
RefSeq	NP_037224.1. NM_013092.1.
UniGene	Rn.10182.

### 3D structure databases

ProteinModelPortal	P50339.
SMR	P50339. Positions 22-247.
ModBase	Search...

### Protein family/group databases

MEROPS	S01.150.
--------	----------

### Proteomic databases

PRIDE	P50339.
-------	---------

### Genome annotation databases

GeneID	25627.
KEGG	rno:25627.

### Organism-specific databases

CTD	25627.
RGD	2365. Cma1.

### Phylogenomic databases

eggNOG	roNOG06568.
HOVERGEN	HBG013304.

### Enzyme and pathway databases

BRENDA	3.4.21.39. 248.
--------	-----------------

### Gene expression databases

Genevestigator	P50339.
----------------	---------

**Family and domain databases**

InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
NextBio	607417.

**Entry information**

Entry name	CMA1_RAT
Accession	Primary (citable) accession number: <b>P50339</b> Secondary accession number(s): Q9R2C8
Entry history	Integrated into    October 1, 1996 UniProtKB/Swiss- Prot: Last sequence    October 1, 1996 update: Last modified:    January 11, 2011 This is version 75 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program

**Relevant documents**

Peptidase families  
Classification of peptidase families and list of entries

SIMILARITY comments  
Index of protein domains and families



**P00784 (PAPA1\_CARPA)** ★ Reviewed, UniProtKB/Swiss-Prot

Last modified January 11, 2011. Version 95.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Papain</b> EC=3.4.22.2 <i>Alternative name(s):</i> Papaya proteinase I Short name=PPI Allergen=Car p 1
Organism	<b>Carica papaya (Papaya)</b>
Taxonomic identifier	3649 [NCBI]
Taxonomic lineage	Eukaryota › Viridiplantae › Streptophyta › Embryophyta › Tracheophyta › Spermatophyta › Magnoliophyta › eudicotyledons › core eudicotyledons › rosids › malvids › Brassicales › Caricaceae › Carica

**Protein attributes**

Sequence length	345 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Catalytic activity	Hydrolysis of proteins with broad specificity for peptide bonds, but preference for an amino acid bearing a large hydrophobic side chain at the P2 position. Does not accept Val in P1'.
Allergenic properties	Causes an allergic reaction in human.
Sequence similarities	Belongs to the peptidase C1 family.

**Ontologies****Keywords**

Disease	Allergen
Domain	Signal
Molecular function	Hydrolase Protease Thiol protease
PTM	Disulfide bond Zymogen

Technical term		3D-structure Direct protein sequencing				
Gene Ontology (GO)						
Biological process		proteolysis Inferred from electronic annotation. Source: InterPro				
Molecular function		cysteine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro				
Complete GO annotation...						
Sequence annotation (Features)						
	Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
Molecule processing						
<input checked="" type="checkbox"/>	Signal peptide	1 – 18	18	Potential		
<input checked="" type="checkbox"/>	Propeptide	19 – 133	115	Activation peptide		PRO_000002640
<input checked="" type="checkbox"/>	Chain	134 – 345	212	Papain		PRO_000002640
Sites						
<input type="checkbox"/>	Active site	158	1	Ref.4		
<input type="checkbox"/>	Active site	292	1			
<input type="checkbox"/>	Active site	308	1	Ref.4		
Amino acid modifications						
<input type="checkbox"/>	Disulfide bond	155 ↔ 196		Ref.2		
<input type="checkbox"/>	Disulfide bond	189 ↔ 228		Ref.2		
<input type="checkbox"/>	Disulfide bond	286 ↔ 333		Ref.2		
Experimental info						
<input type="checkbox"/>	Sequence conflict	180	1	E → Q AA sequence Ref.2		
<input type="checkbox"/>	Sequence conflict	219 – 220	2	YP → PY AA sequence Ref.2		
<input type="checkbox"/>	Sequence conflict	251	1	E → Q AA sequence Ref.2		

<input type="checkbox"/>	Sequence conflict	268	1	E → Q AA sequence Ref.2	
--------------------------	-------------------	-----	---	-------------------------------	--

Secondary structure

1 . . . . . 34

☒ Helix    ☐ Strand    ☒ Turn

Details...

Sequences

Sequence	Length	Mass (Da)
<input type="checkbox"/> P00784 [UniParc]. Last modified January 1, 1988. Version 1. Checksum: 82D9FB35EDCA12EF	FASTA	345 38,922
<div><div>102030405060</div><div>MAMIPSISKL LFVAICLFVY MGLSFGDFS I VGYSQNDLTS TERLIQLFES WMLKHNKIYK</div></div> <div><div>708090100110120</div><div>NIDEKIYRFE IFKDNLKYID ETNKKNSYW LGLNVFADMS NDEFKEKYTG SIAGNYTTTE</div></div> <div><div>130140150160170180</div><div>LSYEEVLNDG DVNIPEYVDW RQKGAVTPVK NQGSCGSCWA FSAVVTIEGI IKIRTGNLNE</div></div> <div><div>190200210220230240</div><div>YSEQELLD CD RRSYGCNGGY PWSALQLVAQ YGIHYRNTYP YEGVQRYCRS REKGPYAAKT</div></div> <div><div>250260270280290300</div><div>DGVRQVQPIN EGALLYSIAN QPVSVVLEAA GKDFQLYRGG IFVGPCGNKV DHAVA AVGYG</div></div> <div><div>310320330340</div><div>PNYILIKNSW GTGWGENGYI RIKRGTGNSY GVCGLYTSSF YPVKN</div></div>		

« Hide

References

[1]	<b>"Cloning and sequencing of papain-encoding cDNA."</b> Cohen L.W., Coghlan V.M., Dihel L.C. Gene 48:219-227(1986) [PubMed: 2881845] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].
[2]	<b>"The complete amino acid sequence of papain. Additions and corrections."</b> Mitchel R.E.J., Chaiken I.M., Smith E.L. J. Biol. Chem. 245:3485-3492(1970) [PubMed: 5470818] [Abstract] <u>Cited for:</u> PROTEIN SEQUENCE OF 134-345.
[3]	<b>"A reinvestigation of residues 64-68 and 175 in papain. Evidence that residues 64 and 175 are asparagine."</b> Husain S.S., Lowe G. Biochem. J. 116:689-692(1970) [PubMed: 5435495] [Abstract] <u>Cited for:</u> SEQUENCE REVISION TO 197.
[4]	<b>"Structure of papain."</b> Drenth J., Jansonius J.N., Koekoek R., Swen H.M., Wolthers B.G.

	Nature 218:929-932(1968) [PubMed: 5681232] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
[5]	<b>"Structure of papain refined at 1.65-A resolution."</b> Kamphuis I.G., Kalk K.H., Swarte M.B.A., Drenth J. J. Mol. Biol. 179:233-256(1984) [PubMed: 6502713] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS).
[6]	<b>"The refined 2.4 A X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction."</b> Stubbs M.T., Laber B., Bode W., Huber R., Jerala R., Lenarcic B., Turk V. EMBO J. 9:1939-1947(1990) [PubMed: 2347312] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
[7]	<b>"Crystal structure of papain-succinyl-Gln-Val-Val-Ala-Ala-p-nitroanilide complex at 1.7-A resolution: noncovalent binding mode of a common sequence of endogenous thiol protease inhibitors."</b> Yamamoto A., Tomoo K., Doi M., Ohishi H., Inoue M., Ishida T., Yamamoto D., Tsuboi S., Okamoto H., Okada Y. Biochemistry 31:11305-11309(1992) [PubMed: 1445868] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
[8]	<b>"Structure of monoclinic papain at 1.60-A resolution."</b> Pickersgill R.W., Harris G.W., Garman E. Acta Crystallogr. B 48:59-67(1992) <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
+	Additional computationally mapped references.

## Web resources

Worthington enzyme manual

## Cross-references

### Sequence databases

<ul style="list-style-type: none"> <li>⊗ EMBL</li> <li>⊗ GenBank</li> <li>⊗ DDBJ</li> </ul>	M15203 mRNA. Translation: AAB02650.1.
---	---------------------------------------

PIR	PPPA. A26466.
-----	---------------

### 3D structure databases

<ul style="list-style-type: none"> <li>⊗ PDBe</li> <li>⊗ RCSB PDB</li> <li>⊗ PDBj</li> </ul>	Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
	1BP4	X-ray	2.20	A	134-345	[«]
	1BQI	X-ray	2.50	A	134-345	[«]
	1CVZ	X-ray	1.70	A	134-345	[«]
	1EFF	model	-	A	134-345	[«]
	1KHP	X-ray	2.00	A	134-345	[«]
	1KHQ	X-ray	1.60	A	134-345	[«]
	1PAD	X-ray	2.80	A	134-345	[«]
	1PE6	X-ray	2.10	A	134-345	[«]
	1PIP	X-ray	1.70	A	134-345	[«]
	1POP	X-ray	2.10	A	134-345	[«]
	1PPD	X-ray	2.00	A	134-345	[«]
	1PPN	X-ray	1.60	A	134-345	[«]
	1PPP	X-ray	1.90	A	134-345	[«]



	1STF X-ray 2.37 E 134-345 [»] 2CIO X-ray 1.50 A 134-345 [»] 2PAD X-ray 2.80 A 134-345 [»] 3E1Z X-ray 1.86 B 134-345 [»] 3IMA X-ray 2.03 A/C 134-345 [»] 3LFY X-ray 2.60 A/C 134-345 [»] 4PAD X-ray 2.80 A 134-345 [»] 5PAD X-ray 2.80 A 134-345 [»] 6PAD X-ray 2.80 A 134-345 [»] 9PAP X-ray 1.65 A 134-345 [»]
ProteinModelPortal	P00784.
SMR	P00784. Positions 38-345.
ModBase	Search...
<b>Protein family/group databases</b>	
Allergome	709. Car p 1.
MEROPS	C01.001.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.22.2. 18730.
<b>Family and domain databases</b>	
InterPro	IPR000169. Pept_cys_AS. IPR013128. Peptidase_C1A. IPR000668. Peptidase_C1A_C. IPR013201. Prot_inhib_I29. [Graphical view]
PANTHER	PTHR12411. Peptidase_C1A. 1 hit.
Pfam	PF08246. Inhibitor_I29. 1 hit. PF00112. Peptidase_C1. 1 hit. [Graphical view]
PRINTS	PR00705. PAPAIN.
SMART	SM00848. Inhibitor_I29. 1 hit. SM00645. Pept_C1. 1 hit. [Graphical view]
PROSITE	PS00640. THIOL_PROTEASE_ASN. 1 hit. PS00139. THIOL_PROTEASE_CYS. 1 hit. PS00639. THIOL_PROTEASE_HIS. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
BindingDB	P00784.
<b>Entry information</b>	
Entry name	PAPA1_CARPA
Accession	Primary (citable) accession number: <b>P00784</b>

Entry history	Integrated into      July 21, 1986 UniProtKB/Swiss- Prot: Last sequence      January 1, 1988 update: Last modified:      January 11, 2011 This is version 95 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Plant Protein Annotation Program
<b>Relevant documents</b>	
Allergens Nomenclature of allergens and list of entries	
PDB cross-references Index of Protein Data Bank (PDB) cross-references	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



**P08246 (ELNE\_HUMAN) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 129.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Neutrophil elastase</b> EC=3.4.21.37 <i>Alternative name(s):</i> Bone marrow serine protease Elastase-2 Human leukocyte elastase Short name=HLE Medullasin PMN elastase
Gene names	Name: <b>ELANE</b> Synonyms: ELA2
Organism	<b>Homo sapiens (Human)</b> [Complete proteome]
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo

**Protein attributes**

Sequence length	267 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Modifies the functions of natural killer cells, monocytes and granulocytes. Inhibits C5a-dependent neutrophil enzyme release and chemotaxis. [Ref.15]
Catalytic activity	Hydrolysis of proteins, including elastin. Preferential cleavage: Val- -Xaa > Ala- -Xaa.
Subunit structure	Interacts with NOTCH2NL. [Ref.16]
Tissue specificity	Bone marrow cells.
Involvement in disease	Defects in ELANE are a cause of cyclic haematopoiesis (CH) [MIM:162800]; also known as cyclic neutropenia. CH is an autosomal dominant disease in which blood-cell production from the bone marrow oscillates with 21-day periodicity. Circulating neutrophils vary between almost normal numbers and zero. During intervals of neutropenia, affected individuals are at risk for opportunistic infection. Monocytes, platelets, lymphocytes and reticulocytes also cycle with the same frequency. [Ref.16] [Ref.21]

## Protein

Translations of Life

Display Settings: GenPept

## neutrophil elastase [Homo sapiens]

GenBank: AAA36359.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAA36359 267 aa linear PRI 27-APR-1993  
DEFINITION neutrophil elastase [Homo sapiens].  
ACCESSION AAA36359  
VERSION AAA36359.1 GI:386981  
DBSOURCE locus HUMNEL1 accession M20199.1  
locus HUMNEL2 accession M20200.1  
locus HUMNEL3 accession M20201.1  
locus HUMNEL4 accession M20202.1  
locus HUMNEL5 accession M20203.1

KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.

REFERENCE 1 (residues 1 to 267)  
AUTHORS Takahashi,H., Nukiwa,T., Yoshimura,K., Quick,C.D., States,D.J.,  
Holmes,M.D., Whang-Peng,J., Knutsen,T. and Crystal,R.G.  
TITLE Structure of the human neutrophil elastase gene  
JOURNAL J. Biol. Chem. 263 (29), 14739-14747 (1988)  
PUBMED 2902087  
COMMENT On Aug 28, 1993 this sequence version replaced gi:189149.  
Method: conceptual translation.

FEATURES  
source Location/Qualifiers  
1..267  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
Protein 1..267  
/name="neutrophil elastase"  
Region 30..245  
/region\_name="Tryp\_SPC"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cd00190"  
/db\_xref="CDD:29152"  
Site 30  
/site\_type="cleavage"  
/db\_xref="CDD:29152"  
Site order(70,117,202)  
/site\_type="active"  
/db\_xref="CDD:29152"  
Site order(196,217,219)  
/site\_type="other"  
/note="substrate binding sites"  
/db\_xref="CDD:29152"  
CDS 1..267  
/coded\_by="join(M20199.1:1445..1511,M20200.1:50..206,  
M20201.1:51..192,M20202.1:49..279,M20203.1:49..255)"

ORIGIN  
1 mtlgrllacl flacvlpall lggtalasei vggrrarpha wpfmvsllqlr gghfcgatli  
61 apnfvmsaah cvanvnrav rvvlgaahls rreptrqvfa vgrifengyd pvnllndivi  
121 lqlngsatin anvgvaqlpa qgrrlgnvgv clangwgllg rnrngiasvlq elnvtvvtsl  
181 crrsnvctlv rgrqagvcfg dsqspvcng lihngiasfvr ggcasglypd afapvaqfvr  
241 widsiiqrse dnpcphprdp dpasrth  
//



## Protein

Translations of Life

Display Settings: GenPept

## elastase/medullasin precursor (EC 3.4.21.37) [Homo sapiens]

GenBank: AAA36173.1

[FASTA](#) [Graphics](#)Go to:

LOCUS AAA36173 267 aa linear PRI 11-JUN-1993  
DEFINITION elastase/medullasin precursor (EC 3.4.21.37) [Homo sapiens].  
ACCESSION AAA36173  
VERSION AAA36173.1 GI:307123  
DBSOURCE locus HUMLEUELA accession M34379.1  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 267)  
AUTHORS Okano,K., Aoki,Y., Shimizu,H. and Naruto,M.  
TITLE Functional expression of human leukocyte elastase (HLE)/medullasin  
in eukaryotic cells  
JOURNAL Biochem. Biophys. Res. Commun. 167 (3), 1326-1332 (1990)  
PUBMED 2322278  
COMMENT On Jul 26, 1993 this sequence version replaced gi:187117.  
Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..267  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
Protein 1..267  
/name="elastase/medullasin precursor (EC 3.4.21.37)"  
sig\_peptide 1..27  
/note="elastase/medullasin signal peptide"  
mat\_peptide 30..267  
/product="elastase/medullasin"  
Region 30..245  
/region\_name="Tryp\_Spc"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cd00190"  
/db\_xref="CDD:29152"  
Site 30  
/site\_type="cleavage"  
/db\_xref="CDD:29152"  
Site order(70,117,202)  
/site\_type="active"  
/db\_xref="CDD:29152"  
Site order(196,217,219)  
/site\_type="other"  
/note="substrate binding sites"  
/db\_xref="CDD:29152"  
CDS 1..267  
/coded\_by="M34379.1:39..842"  
ORIGIN  
1 mtlgrriac1 flacvlpall lggatalasei vggrrarpha wpmvslqlr gghfcgatli  
61 apnfvmsaah cvanvnvrvav rvvlgaahls rreptrqvfa vqrifengyd pvnllndivi  
121 lqlngsatin anvgvaqlpa qgrrlgnvgvq clamgwllg rnrngiasvlq elnvtvvtsl  
181 crrenvctlv rgrqagvcfg dsgsplvcng lihngiasfvr ggcasglypd afapvaqfvn  
241 widsiiqrse dnpcphprdp dpasrth  
//

**P37357 (EL2A\_HORSE) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified January 11, 2011. Version 52.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Neutrophil elastase 2A</b> EC=3.4.21.- <i>Alternative name(s):</i> Proteinase 2A
Organism	<b><i>Equus caballus</i> (Horse)</b>
Taxonomic identifier	9796 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Laurasiatheria › Perissodactyla › Equidae › Equus

**Protein attributes**

Sequence length	85 AA.
Sequence status	Fragments.
Protein existence	Evidence at protein level.









**General annotation (Comments)**

Function	May be involved in the degradation of connective tissue in chronic lung disease.
Sequence similarities	Belongs to the peptidase S1 family. Elastase subfamily. Contains 1 peptidase S1 domain.


**Ontologies**

<b>Keywords</b>	
Molecular function	Hydrolase Protease Serine protease
Technical term	Direct protein sequencing
<b>Gene Ontology (GO)</b>	
Biological process	proteolysis Inferred from electronic annotation. Source: InterPro
Molecular function	serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro
Complete GO annotation...	

Sequence annotation (Features)

Feature key	Position (s)	Length	Description	Graphical view	Feature identifier
Molecule processing					
 Chain	1 – 85	85	Neutrophil elastase 2A		PRO_00000886
Regions					
 Domain	1 – 85	85	Peptidase S1		
Sites					
 Active site	67	1	Charge relay system		
Experimental info					
 Non-adjacent residues	34 – 35	2			
 Non-adjacent residues	59 – 60	2			

Sequences

Sequence	Length	Mass (Da)
 P37357 [UniParc]. Last modified October 1, 1994. Version 1. Checksum: 0ABE7D44A1B90E26	FASTA 85	8,893
<div><div><div>102030405060</div><div>IVGGRAAEPI SRPYMVSLQI RGNPGSHFCG GTLIMGWGRL GTREPLPXVL QELNVTVVTA</div><div>7080</div><div>GICFGDSGGP LICNGVAQGV FSFVR</div></div><div>« Hide</div></div>		

References

[1] "Structural and functional characterization of elastases from horse neutrophils."  
Dubin A., Potempa J., Travis J.  
Biochem. J. 300:401-406(1994) [PubMed: 7516152] [Abstract]  
Cited for: PROTEIN SEQUENCE.  
Tissue: Neutrophil.

**Cross-references****Sequence databases**

PIR	S44461.
-----	---------

**3D structure databases**

ProteinModelPortal	P37357.
--------------------	---------

SMR	P37357. Positions 1-34.
-----	-------------------------

ModBase	Search...
---------	-----------

**Protein family/group databases**

MEROPS	S01.131.
--------	----------

**Family and domain databases**

InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. [Graphical view]
----------	--

Pfam	PF00089. Trypsin. 2 hits. [Graphical view]
------	---

SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
-------	---

SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
--------	--------------------------------

PROSITE	PS50240. TRYPSIN_DOM. Partial match. PS00134. TRYPSIN_HIS. Partial match. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
---------	--

ProtoNet	Search...
----------	-----------

**Entry information**

Entry name	EL2A_HORSE
------------	------------

Accession	Primary (citable) accession number: <b>P37357</b>
-----------	---

Entry history	Integrated into      October 1, 1994 UniProtKB/Swiss- Prot: Last sequence      October 1, 1994 update: Last modified:      January 11, 2011 This is version 52 of the entry and version 1 of the sequence. [Complete history]
---------------	--

Entry status	Reviewed (UniProtKB/Swiss-Prot)
--------------	---------------------------------

Annotation program	Chordata Protein Annotation Program
--------------------	-------------------------------------

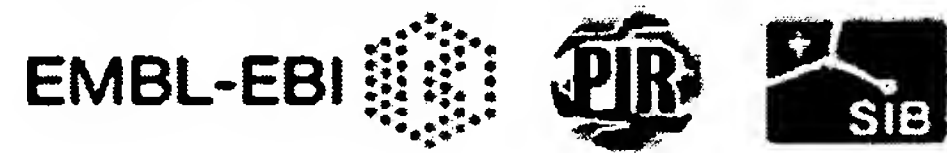
**Relevant documents**

Peptidase families Classification of peptidase families and list of entries
--



SIMILARITY comments  
Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



**P37358 (EL2B\_HORSE) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified January 11, 2011. Version 50.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Neutrophil elastase 2B</b> EC=3.4.21.- <i>Alternative name(s):</i> Proteinase 2B
Organism	<b>Equus caballus (Horse)</b>
Taxonomic identifier	9796 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Laurasiatheria › Perissodactyla › Equidae › Equus

**Protein attributes**

Sequence length	73 AA.
Sequence status	Fragments.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	May be involved in the degradation of connective tissue in chronic lung disease.
Sequence similarities	Belongs to the peptidase S1 family. Elastase subfamily. Contains 1 peptidase S1 domain.

**Ontologies****Keywords**

Molecular function	Hydrolase Protease Serine protease
Technical term	Direct protein sequencing

**Gene Ontology (GO)**

Biological process	proteolysis Inferred from electronic annotation. Source: InterPro
Molecular function	serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro

Complete GO annotation...

## Sequence annotation (Features)

	Feature key	Position (s)	Length	Description	Graphical view	Feature identifier
<b>Molecule processing</b>						
<input type="checkbox"/>	Chain	1 – 73	73	Neutrophil elastase 2B		PRO_00000886
<b>Regions</b>						
<input type="checkbox"/>	Domain	1 – 73	73	Peptidase S1		
<b>Sites</b>						
<input type="checkbox"/>	Active site	64	1	Charge relay system	I	
<b>Experimental info</b>						
<input type="checkbox"/>	Non-adjacent residues	31 – 32	2			
<input type="checkbox"/>	Non-adjacent residues	56 – 57	2			

## Sequences

	Sequence	Length	Mass (Da)
<input type="checkbox"/>	P37358 [UniParc]. Last modified October 1, 1994. Version 1. Checksum: AFFB0B330DB69041	FASTA 73	7,615
<div> <div> <div>10</div> <div>20</div> <div>30</div> <div>40</div> <div>50</div> <div>60</div> </div> <div> <div>IVGGRPARPH</div> <div>AWPFMASLQR</div> <div>RGGHFCGATL</div> <div>IMGWGQLGTN</div> <div>RPLPSVLQEL</div> <div>NVTVVTAGIC</div> </div> </div> <div> <div>70</div> <div>FGDSGGPLVC</div> <div>NGL</div> </div>			
« Hide			

## References

- [1] "Structural and functional characterization of elastases from horse neutrophils."  
Dubin A., Potempa J., Travis J.  
Biochem. J. 300:401-406(1994) [PubMed: 7516152] [Abstract]  
Cited for: PROTEIN SEQUENCE.  
Tissue: Neutrophil.

**Cross-references****Sequence databases**

PIR	S44462.
-----	---------

**3D structure databases**

ProteinModelPortal	P37358.
--------------------	---------

SMR	P37358. Positions 1-73.
-----	-------------------------

ModBase	Search...
---------	-----------

**Protein family/group databases**

MEROPS	S01.131.
--------	----------

**Family and domain databases**

InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. [Graphical view]
----------	--

Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
------	--

SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
--------	--------------------------------

PROSITE	PS50240. TRYPSIN_DOM. Partial match. PS00134. TRYPSIN_HIS. Partial match. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
---------	--

ProtoNet	Search...
----------	-----------

**Entry information**

Entry name	EL2B_HORSE
------------	------------

Accession	Primary (citable) accession number: <b>P37358</b>
-----------	---

Entry history	Integrated into      October 1, 1994 UniProtKB/Swiss- Prot: Last sequence      October 1, 1994 update: Last modified:      January 11, 2011 This is version 50 of the entry and version 1 of the sequence. [Complete history]
---------------	--

Entry status	Reviewed (UniProtKB/Swiss-Prot)
--------------	---------------------------------

Annotation program	Chordata Protein Annotation Program
--------------------	-------------------------------------

**Relevant documents**

Peptidase families



Classification of peptidase families and list of entries

SIMILARITY comments

Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



**Q9Z284 (Q9Z284\_MOUSE) ★ Unreviewed, UniProtKB/TrEMBL**

Last modified April 5, 2011. Version 77.

**Names and origin**

Protein names	<i>Submitted name:</i> Neutrophil elastase <a href="#">EMBL AAC79702.1</a>
Gene names	Name: <b>Elane</b> <a href="#">MGI 2679229</a> Synonyms: <b>Cela2a</b> <a href="#">MGI 95316</a> , <b>Ela2</b> <a href="#">MGI 95316</a> <a href="#">MGI 2679229</a> , <b>Ela2a</b> <a href="#">MGI 95316</a>
Organism	<b>Mus musculus (Mouse)</b> <a href="#">EMBL AAC79702.1</a>
Taxonomic identifier	10090 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires › Rodentia › Sciurognathi › Muroidea › Muridae › Murinae › Mus › Mus

**Protein attributes**

Sequence length	207 AA.
Sequence status	Complete.
Protein existence	Evidence at transcript level.

**General annotation (Comments)**

Sequence similarities	Belongs to the peptidase S1 family. <a href="#">RuleBase RU000360V4</a>
-----------------------	---

**Ontologies****Gene Ontology (GO)**

Biological process	acute inflammatory response to antigenic stimulus Inferred from direct assay. Source: MGI  leukocyte migration Inferred from mutant phenotype. Source: MGI  negative regulation of growth of symbiont in host Inferred from genetic interaction. Source: MGI  neutrophil mediated killing of fungus Inferred from mutant phenotype. Source: MGI  phagocytosis Inferred from mutant phenotype. Source: MGI  positive regulation of immune response Inferred from genetic interaction. Source: MGI  proteolysis Inferred from electronic annotation. Source: InterPro  response to lipopolysaccharide
--------------------	--

	Inferred from genetic interaction. Source: MGI response to yeast Inferred from mutant phenotype. Source: MGI
Cellular component	cytoplasm Inferred from direct assay. Source: MGI
Molecular function	serine-type endopeptidase activity Inferred from direct assay. Source: MGI
Complete GO annotation...	

Sequences

Sequence	Length	Mass (Da)
<div><div><div><div><div><div></div><div>Q9Z284 [UniParc].</div><div>Last modified May 1, 1999. Version 1.</div><div>Checksum: 1F4E834DDD173DEE</div></div></div><div><div><div>102030405060</div><div>MALGRLSSRT LAAMLLALFL GGPALPSEIV GGRPARPHAW PFMASLQRRG GHFCGATLIA</div><div>708090100110120</div><div>RNFVMSAVHC VTGNFRSVQV VLGADLRRQ ERTRQTFSVQ RIFENGFDPS QLLNDIVIIQ</div><div>130140150160170180</div><div>LNGSATINAN VQVAQLPAQG QGVGDRTPCL AMGWGRLGTN RPSPSVLQEL NVTVTNMCP</div><div>190200</div><div>RRVNVCTLVP RRQAGICFVS TLCRRRV</div></div></div></div></div></div>	FASTA	20722,526
« Hide		

References

[1]

"Characterization and localization of the genes for mouse proteinase-3 and neutrophil elastase."  
Sturrock A., Franklin K.F., Wu S.Q., Hoidal J.R.  
Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases  
Cited for: NUCLEOTIDE SEQUENCE.  
Strain: 129/SvJ EMBL AAC79702.1

+

Additional computationally mapped references.

Cross-references

Sequence databases	
<div><div><div><div><div></div><div>EMBL</div></div><div><div></div><div>GenBank</div></div><div><div></div><div>DDBJ</div></div></div></div></div>	AF082186 Genomic DNA. Translation: AAC79702.1.
IPI	IPI00785483.
UniGene	Mm.262194.

<b>3D structure databases</b>	
HSSP	HSSP built from PDB template 1PPF based on UniProtKB P08246.
ProteinModelPortal	Q9Z284.
SMR	Q9Z284. Positions 29-198.
ModBase	Search...
<b>Protein-protein interaction databases</b>	
STRING	Q9Z284.
<b>Organism-specific databases</b>	
MGI	MGI:2679229. Elane.
<b>Phylogenomic databases</b>	
GeneTree	ENSGT00580000081367.
HOVERGEN	HBG013304.
InParanoid	Q9Z284.
<b>Gene expression databases</b>	
ArrayExpress	Q9Z284.
Bgee	Q9Z284.
Genevestigator	Q9Z284.
<b>Family and domain databases</b>	
InterPro	IPR009003. Pept_cys/ser_Trypsin-like. IPR001254. Peptidase_S1_S6. [Graphical view]
Pfam	PF00089. Trypsin. 1 hit. [Graphical view]
SMART	SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS50240. TRYPSIN_DOM. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
SOURCE	Search...
<b>Entry information</b>	
Entry name	Q9Z284_MOUSE
Accession	Primary (citable) accession number: <b>Q9Z284</b>



Entry history	Integrated into May 1, 1999 UniProtKB/TrEMBL: Last sequence May 1, 1999 update: Last modified: April 5, 2011 This is version 77 of the entry and version 1 of the sequence. [Complete history]
Entry status	Unreviewed (UniProtKB/TrEMBL)

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



**P48740 (MASP1\_HUMAN) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 116.

**Names and origin**

Protein names	<p><i>Recommended name:</i> <b>Mannan-binding lectin serine protease 1</b> EC=3.4.21.-</p> <p><i>Alternative name(s):</i> Complement factor MASP-3 Complement-activating component of Ra-reactive factor Mannose-binding lectin-associated serine protease 1 Short name=MASP-1 Mannose-binding protein-associated serine protease Ra-reactive factor serine protease p100 Short name=RaRF Serine protease 5</p> <p><u>Cleaved into the following 2 chains:</u></p> <ol style="list-style-type: none"><li>1. <b>Mannan-binding lectin serine protease 1 heavy chain</b></li><li>2. <b>Mannan-binding lectin serine protease 1 light chain</b></li></ol>
Gene names	<p>Name: <b>MASP1</b></p> <p>Synonyms: CRARF, CRARF1, PRSS5</p>
Organism	<b>Homo sapiens (Human) [Complete proteome]</b>
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo

**Protein attributes**

Sequence length	699 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Functions in the lectin pathway of complement, which performs a key role in innate immunity by recognizing pathogens through patterns of sugar moieties and neutralizing them. The lectin pathway is triggered upon binding of mannan-binding lectin (MBL) and ficolins to sugar moieties which leads to activation of the associated proteases MASP1 and MASP2. Functions as an endopeptidase and may activate MASP2 or C2 or directly activate C3 the key component of complement reaction. Isoform 2 may have an inhibitory effect on the activation of the
----------	--

	lectin pathway of complement or may cleave IGFBP5. <a href="#">Ref.5</a> <a href="#">Ref.24</a>
Enzyme regulation	Inhibited by SERPING1 and A2M. <a href="#">Ref.17</a> <a href="#">Ref.20</a>
Subunit structure	Homodimer. Interacts with the oligomeric lectins MBL2, FCN2 and FCN3; triggers the lectin pathway of complement through activation of C3. Interacts with SERPING1. <a href="#">Ref.11</a> <a href="#">Ref.13</a> <a href="#">Ref.14</a> <a href="#">Ref.15</a> <a href="#">Ref.16</a> <a href="#">Ref.18</a> <a href="#">Ref.19</a> <a href="#">Ref.25</a>
Subcellular location	Secreted <a href="#">Ref.5</a>
Tissue specificity	Protein of the plasma which is primarily expressed by liver. <a href="#">Ref.5</a> <a href="#">Ref.1</a> <a href="#">Ref.2</a> <a href="#">Ref.12</a>
Post-translational modification	The iron and 2-oxoglutarate dependent 3-hydroxylation of aspartate and asparagine is (R) stereospecific within EGF domains <a href="#">By similarity</a> .  N-glycosylated. Some N-linked glycan are of the complex-type <a href="#">By similarity</a> . <a href="#">Ref.11</a> <a href="#">Ref.25</a> <a href="#">Ref.22</a>  Autoproteolytic processing of the proenzyme produces the active enzyme composed on the heavy and the light chain held together by a disulfide bond. Isoform 1 but not isoform 2 is activated through autoproteolytic processing.
Sequence similarities	Belongs to the peptidase S1 family. Contains 2 CUB domains. Contains 1 EGF-like domain. Contains 1 peptidase S1 domain. Contains 2 Sushi (CCP/SCR) domains.
Biophysicochemical properties	<u>Kinetic parameters:</u> K <sub>M</sub> =0.10 mM for Ac-Gly-Lys-OMe (at 30 degrees Celsius) <a href="#">Ref.17</a> <a href="#">Ref.20</a>  K <sub>M</sub> =310 µM for Bz-Arg-OEt (at 30 degrees Celsius)  K <sub>M</sub> =4.8 µM for C2 (at 37 degrees Celsius)
Sequence caution	The sequence AAH39724.1 differs from that shown. Reason: Erroneous initiation.
<b>Ontologies</b>	
<b>Keywords</b>	
Biological process	Complement activation lectin pathway Immunity Innate immunity
Cellular component	Secreted
Coding sequence diversity	Alternative splicing Polymorphism
Domain	EGF-like domain Repeat Signal Sushi

Ligand	Calcium Metal-binding
Molecular function	Hydrolase Protease Serine protease
PTM	Autocatalytic cleavage Disulfide bond Glycoprotein Hydroxylation
Technical term	3D-structure Complete proteome Direct protein sequencing
<b>Gene Ontology (GO)</b> Biological process	complement activation, lectin pathway Inferred from mutant phenotype <a href="#">Ref.24</a> . Source: UniProtKB  negative regulation of complement activation Inferred from direct assay <a href="#">Ref.5</a> . Source: UniProtKB  proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	extracellular space Inferred from direct assay <a href="#">Ref.5</a> . Source: UniProtKB
Molecular function	calcium ion binding Inferred from direct assay <a href="#">Ref.25</a> . Source: UniProtKB  calcium-dependent protein binding Inferred from physical interaction <a href="#">Ref.19</a> <a href="#">Ref.19</a> <a href="#">Ref.25</a> . Source: UniProtKB  protein binding Inferred from physical interaction <a href="#">Ref.16</a> <a href="#">Ref.5</a> <a href="#">Ref.18</a> . Source: UniProtKB  protein homodimerization activity Inferred from physical interaction <a href="#">Ref.25</a> . Source: UniProtKB  serine-type endopeptidase activity Inferred from direct assay <a href="#">Ref.17</a> . Source: UniProtKB
Complete GO annotation...	
<b>Alternative products</b>	
This entry describes 4 isoforms produced by <b>alternative splicing</b> . <a href="#">[Align]</a> <a href="#">[Select]</a>	
<b>Isoform 1 (identifier: P48740-1)</b>	
<i>This isoform has been chosen as the 'canonical' sequence. All positional information in this entry refers to it. This is also the sequence that appears in the downloadable versions of the entry.</i>	
<b>Isoform 2 (identifier: P48740-2)</b>	
Also known as: MASP-3;	



The sequence of this isoform differs from the canonical sequence as follows: ·  
435-435: V → ECGQPSRSLP...QSVVEPQVER  
436-699: Missing.

Note: Glycosylated on Asn-533 and Asn-599.














Isoform 3 (identifier: P48740-3)

The sequence of this isoform differs from the canonical sequence as follows:  
364-380: IVDCRAPGELEHGLITF → KNEIDLESELKSEQVTE  
381-699: Missing.

Isoform 4 (identifier: P48740-4)

The sequence of this isoform differs from the canonical sequence as follows:  
1-113: Missing.  
435-435: V → ECGQPSRSLP...QSVVEPQVER  
436-699: Missing.

Sequence annotation (Features)

	Feature key	Position(s)	Length	Description	Graphical view
Molecule processing					
	Signal peptide	1 – 19	19	Ref.11	
	Chain	20 – 699	680	Mannan-binding lectin serine protease 1	
	Chain	20 – 448	429	Mannan-binding lectin serine protease 1 heavy chain	
	Chain	449 – 699	251	Mannan-binding lectin serine protease 1 light chain	
Regions					
	Domain	20 – 138	119	CUB 1	
	Domain	139 – 182	44	EGF-like; calcium-binding	
	Domain	185 – 297	113	CUB 2	
	Domain	299 – 364	66	Sushi 1	
	Domain	365 – 434	70	Sushi 2	
	Domain	449 – 696	248	Peptidase S1	
	Region	20 – 278	259	Interaction with FCN2	
	Region	20 – 184	165	Homodimerization By similarity	
	Region	20 – 184	165	Interaction with MBL2	
Sites					
<input type="checkbox"/>	Active site	490	1	Charge relay system By similarity	
<input type="checkbox"/>	Active site	552	1	Charge relay system By similarity	
<input type="checkbox"/>	Active site	646	1	Charge relay system By similarity	

<input type="checkbox"/>	Metal binding	68	1	Calcium 1	I
<input type="checkbox"/>	Metal binding	76	1	Calcium 1	I
<input type="checkbox"/>	Metal binding	121	1	Calcium 1	I
<input type="checkbox"/>	Metal binding	123	1	Calcium 1; via carbonyl oxygen	I
<input type="checkbox"/>	Metal binding	139	1	Calcium 2	I
<input type="checkbox"/>	Metal binding	140	1	Calcium 2; via carbonyl oxygen	I
<input type="checkbox"/>	Metal binding	142	1	Calcium 2	I
<input type="checkbox"/>	Metal binding	159	1	Calcium 2	I
<input type="checkbox"/>	Metal binding	160	1	Calcium 2; via carbonyl oxygen	I
<input type="checkbox"/>	Metal binding	163	1	Calcium 2; via carbonyl oxygen	I
<input type="checkbox"/>	Metal binding	235	1	Calcium 3	I
<input type="checkbox"/>	Metal binding	245	1	Calcium 3	I
<input type="checkbox"/>	Metal binding	282	1	Calcium 3	I
<input type="checkbox"/>	Metal binding	284	1	Calcium 3; via carbonyl oxygen	I
<input type="checkbox"/>	Site	448 – 449	2	Cleavage; by autolysis	I
<b>Amino acid modifications</b>					
<input type="checkbox"/>	Modified residue	159	1	(3R)-3-hydroxyasparagine Potential	I
<input type="checkbox"/>	Glycosylation	49	1	N-linked (GlcNAc...) Ref.22	I
<input type="checkbox"/>	Glycosylation	178	1	N-linked (GlcNAc...) Ref.25 Ref.22	I
<input type="checkbox"/>	Glycosylation	385	1	N-linked (GlcNAc...) Ref.22	I
<input type="checkbox"/>	Glycosylation	407	1	N-linked (GlcNAc...) Ref.22	I
<input type="checkbox"/>	Disulfide bond	73 ↔ 91		Ref.25	II
<input type="checkbox"/>	Disulfide bond	143 ↔ 157		Ref.25	II
<input type="checkbox"/>	Disulfide bond	153 ↔ 166		Ref.25	II
<input type="checkbox"/>	Disulfide bond	168 ↔ 181		Ref.25	II
<input type="checkbox"/>	Disulfide bond	185 ↔ 212		Ref.25	II
<input type="checkbox"/>	Disulfide bond	242 ↔ 260		Ref.25	II
<input type="checkbox"/>	Disulfide bond	301 ↔ 349		By similarity	II
<input type="checkbox"/>	Disulfide bond	329 ↔ 362		By similarity	II
<input type="checkbox"/>	Disulfide bond	367 ↔ 414		By similarity	II
<input type="checkbox"/>	Disulfide bond	397 ↔ 432		By similarity	II
<input type="checkbox"/>	Disulfide bond	436 ↔ 572		Interchain (between heavy and light chains) Potential	I
<input type="checkbox"/>	Disulfide bond	475 ↔ 491		By similarity	II
<input type="checkbox"/>	Disulfide bond	614 ↔ 631		By similarity	
<input type="checkbox"/>	Disulfide bond	642 ↔ 672		By similarity	

Natural variations					
<input checked="" type="checkbox"/>	Alternative sequence	1 – 113	113	Missing in isoform 4.	I
<input checked="" type="checkbox"/>	Alternative sequence	364 – 380	17	IVDCR...GLITF → KNEIDLESELKSEQVTE in isoform 3.	II
<input checked="" type="checkbox"/>	Alternative sequence	381 – 699	319	Missing in isoform 3.	I
<input type="checkbox"/>	Alternative sequence	435	1	V → ECGQPSRSLPSLVKRIIGGR NAEPGLFPWQALIVVEDTSR VPNDKWFGSGALLSASWILT AAHVLRSQRRDTTVIPVSKE HVTVYLGLHDVRDKSGAVNS SAARVVLHPDFNIQNYNNDI ALVQLQEPVPLGPHVMPVCL PRLEPEGPAPHMLGLVAGWG ISNPNTVDIISGTRTLS DVLQYVKLPVVPFAECKTSY ESRSGNYSVTENMFCAGYYE GGKDTCLGDSGGAFVIFDDL SQRWVQGLVSWGGPEECGS KQVYGVYTKVSNYVDWVWEQ MGLPQSVPEPQVER in isoform 2 and isoform 4.	I
<input checked="" type="checkbox"/>	Alternative sequence	436 – 699	264	Missing in isoform 2 and isoform 4.	I
<input type="checkbox"/>	Natural variant	21	1	T → I. [dbSNP:rs1062049]	I
<input type="checkbox"/>	Natural variant	568	1	V → A. [dbSNP:rs13322090]	
<input type="checkbox"/>	Natural variant	679	1	G → R. [dbSNP:rs3774266]	
Experimental info					
<input type="checkbox"/>	Mutagenesis	68	1	E → A or Q: Partial loss of interaction with FCN2, FCN3 and MBL2. [Ref.25]	I
<input type="checkbox"/>	Mutagenesis	77	1	Y → A: Partial loss of interaction with FCN2, FCN3 and MBL2. [Ref.25]	I
<input type="checkbox"/>	Mutagenesis	99	1	E → A: Partial loss of interaction with FCN2, FCN3 and MBL2. [Ref.25]	I
<input type="checkbox"/>	Mutagenesis	121	1	D → A or N: Loss of interaction with FNC2 and FCN3 and partial loss of interaction with MBL2. [Ref.25]	I
<input type="checkbox"/>	Mutagenesis	122	1	F → A: Partial loss of interaction with FCN2, FCN3 and MBL2. [Ref.25]	I
<input type="checkbox"/>	Mutagenesis	123	1	S → A: Partial loss of interaction with FCN2, FCN3 and MBL2. [Ref.25]	I

<input type="checkbox"/>	Mutagenesis	125	1	E → A: Partial loss of interaction with FCN2, FCN3 and MBL2. <a href="#">Ref.25</a>	
<input type="checkbox"/>	Mutagenesis	237	1	H → A: Loss of interaction with FCN2, FCN3 and MBL2. <a href="#">Ref.25</a>	
<input type="checkbox"/>	Mutagenesis	239	1	E → A: Partial loss of interaction with FCN2, FCN3 and MBL2. <a href="#">Ref.25</a>	
<input type="checkbox"/>	Mutagenesis	244	1	Y → A: Loss of interaction with FCN2, FCN3 and MBL2. <a href="#">Ref.25</a>	
<input type="checkbox"/>	Mutagenesis	262	1	E → A: Partial loss of interaction with FCN2, FCN3 and MBL2. <a href="#">Ref.25</a>	
<input type="checkbox"/>	Mutagenesis	274	1	S → A: Partial loss of interaction with FCN2 and FCN3. No effect on interaction with MBL2. <a href="#">Ref.25</a>	
<input type="checkbox"/>	Mutagenesis	283	1	N → A: Partial loss of interaction with FCN2, FCN3 and MBL2. <a href="#">Ref.25</a>	
<input type="checkbox"/>	Mutagenesis	286	1	E → A: Partial loss of interaction with FCN2, FCN3 and MBL2. <a href="#">Ref.25</a>	
<input type="checkbox"/>	Mutagenesis	646	1	S → A: No autoproteolytic processing. <a href="#">Ref.21</a>	
<input type="checkbox"/>	Sequence conflict	2	1	R → K in BAF84375. <a href="#">Ref.6</a>	
<input type="checkbox"/>	Sequence conflict	89	1	T → A in CAH18409. <a href="#">Ref.7</a>	
<input type="checkbox"/>	Sequence conflict	232	1	F → L in BAF84375. <a href="#">Ref.6</a>	
<input type="checkbox"/>	Sequence conflict	235	1	E → Q in BAA05928. <a href="#">Ref.2</a>	
<input type="checkbox"/>	Sequence conflict	235	1	E → Q in BAA34864. <a href="#">Ref.3</a>	
<input type="checkbox"/>	Sequence conflict	285	1	G → A in BAA05928. <a href="#">Ref.2</a>	
<input type="checkbox"/>	Sequence conflict	285	1	G → A in BAA34864. <a href="#">Ref.3</a>	
<input type="checkbox"/>	Sequence conflict	285	1	G → A in BAA89206. <a href="#">Ref.4</a>	
<input type="checkbox"/>	Sequence conflict	392	1	E → G in BAF83846. <a href="#">Ref.6</a>	
<input type="checkbox"/>	Sequence conflict	499	1	E → G in BAA05928. <a href="#">Ref.2</a>	
<input type="checkbox"/>	Sequence conflict	499	1	E → K in BAA04477. <a href="#">Ref.1</a>	
<input type="checkbox"/>	Sequence conflict	499	1	E → K in BAA89206. <a href="#">Ref.4</a>	
<input type="checkbox"/>	Sequence conflict	527	1	D → A in BAA34864. <a href="#">Ref.3</a>	
<input type="checkbox"/>	Sequence conflict	543	1	Q → K in BAA04477. <a href="#">Ref.1</a>	
<input type="checkbox"/>	Sequence conflict	552	1	D → V in BAA34864. <a href="#">Ref.3</a>	
<input type="checkbox"/>	Sequence conflict	643	1	A → S in BAA04477. <a href="#">Ref.1</a>	
<b>Secondary structure</b>					
1.	.....				



■ Helix

■ Strand

■ Turn



Details...


Sequences

Sequence	Length	Mass (Da)
<div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div>Isoform 1 [UniParc].</div><div>Last modified December 16, 2008. Version 3.</div><div>Checksum: 5B37C7FB9F51FD1D</div></div></div><div><div><div>102030405060</div><div>MRWLLLYALCFSLSKASAH TVELNNMFGQ IQSPGYPDSY PSDSEVTWNI TVPDGFRIKL</div></div><div><div>708090100110120</div><div>YFMHFNLESS YLCEYDYVKV ETEDQVLATF CGRETTDTEQ TPGQEVVLSP GSFMSITFRS</div></div><div><div>130140150160170180</div><div>DFSNEERFTG FDAHYMAVDV DECKEREDEE LSCDHYCHNY IGGYYCSCRF GYILHTDNRT</div></div><div><div>190200210220230240</div><div>CRVECSDNLF TQRTGVITSP DFPNPYPKSS ECLYTIELEE GFMVNLQFED IFDIEDHPEV</div></div><div><div>250260270280290300</div><div>PCPYDIYIKI VGPKVLGPFC GEKAPEPIST QSHSVLILFH SDNSGENRGW RLSYRAAGNE</div></div><div><div>310320330340350360</div><div>CPELQPPVHG KIEPSQAKYF FKDQVLVSCD TGYKVLKDNV EMDTFQIECL KDGTWSNKIP</div></div><div><div>370380390400410420</div><div>TCKIVDCRAP GELEHGLITF STRNNLTYYK SEIKYSCQEP YYKMLNNNTG IYTCSAQGVW</div></div><div><div>430440450460470480</div><div>MNKVLGRSLP TCLPVCGLPK FSRKLMARIF NGRPAQKGTI PWIAMLSHLN GQPF CGGSLL</div></div><div><div>490500510520530540</div><div>GSSWIVTAAH CLHQSLDPED PTLRSDLLS PSDFKIILGK HWRLRSDENE QHLGVKHTTL</div></div><div><div>550560570580590600</div><div>HPQYDPNTFE NDVALVELLE SPVLNAFVMP ICLPEGPQQE GAMVIVSGWG KQFLQRFPET</div></div><div><div>610620630640650660</div><div>LMEIEIPIVD HSTCQKAYAP LKKKVTRDMI CAGEKEGGKD ACAGDSGGPM VTLNRERGQW</div></div><div><div>670680690</div><div>YLVGTVSWGD DCGKKDRYGV YSYIHHNKDW IQRVTGVRN</div></div></div></div></div> <div>FASTA69979,247</div>		
<div><div><div></div><div></div></div><div><div><div>Isoform 2 (MASP-3).</div><div>Checksum: 09B5297A6C14283A</div><div>Show »</div></div></div></div> <div>FASTA72881,860</div>		
<div><div><div></div><div></div></div><div><div><div>Isoform 3.</div><div>Checksum: DDED114311A62714</div><div>Show »</div></div></div></div> <div>FASTA38043,640</div>		
<div><div><div></div><div></div></div><div><div><div>Isoform 4.</div><div>Checksum: EEE63886709340FA</div><div>Show »</div></div></div></div> <div>FASTA61568,918</div>		

References

« Hide 'large scale' references

[1]	<p><b>"A new member of the C1s family of complement proteins found in a bactericidal factor, Ra-reactive factor, in human serum."</b>          Takada F., Takayama Y., Hatsuse H., Kawakami M.          Biochem. Biophys. Res. Commun. 196:1003-1009(1993) [PubMed: 8240317] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), TISSUE SPECIFICITY.  <u>Tissue:</u> Liver.</p>
[2]	<p><b>"Molecular characterization of a novel serine protease involved in activation of the complement system by mannose-binding protein."</b>          Sato T., Endo Y., Matsushita M., Fujita T.          Int. Immunol. 6:665-669(1994) [PubMed: 8018603] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), TISSUE SPECIFICITY.  <u>Tissue:</u> Fetal liver.</p>
[3]	<p><b>"Exon structure of the gene encoding the human mannose-binding protein-associated serine protease light chain: comparison with complement C1r and C1s genes."</b>          Endo Y., Sato T., Matsushita M., Fujita T.          Int. Immunol. 8:1355-1358(1996) [PubMed: 8921412] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA].  <u>Tissue:</u> Placenta.</p>
[4]	<p><b>"Gene structure of the P100 serine-protease component of the human Ra-reactive factor."</b>          Takayama Y., Takada F., Nowatari M., Kawakami M., Matsu-ura N.          Mol. Immunol. 36:505-514(1999) [PubMed: 10475605] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA].</p>
[5]	<p><b>"MASP-3 and its association with distinct complexes of the mannan-binding lectin complement activation pathway."</b>          Dahl M.R., Thiel S., Matsushita M., Fujita T., Willis A.C., Christensen T., Vorup-Jensen T., Jensenius J.C.          Immunity 15:127-135(2001) [PubMed: 11485744] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 2), PROTEIN SEQUENCE OF 450-474; 506-526; 539-555; 577-590; 613-621 AND 679-695 (ISOFORM 2), FUNCTION, SUBCELLULAR LOCATION, TISSUE SPECIFICITY.  <u>Tissue:</u> Liver.</p>
[6]	<p><b>"Complete sequencing and characterization of 21,243 full-length human cDNAs."</b>          Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.  Sugano S.          Nat. Genet. 36:40-45(2004) [PubMed: 14702039] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 2 AND 4).  <u>Tissue:</u> Placenta, Teratocarcinoma and Trachea.</p>
[7]	<p><b>"The full-ORF clone resource of the German cDNA consortium."</b>          Bechtel S., Rosenfelder H., Duda A., Schmidt C.P., Ernst U., Wellenreuther R., Mehrle A., Schuster C., Bahr A., Bloecker H., Heubner D., Hoerlein A., Michel G., Wedler H., Koehrer K., Ottenwaelder B., Poustka A., Wiemann S., Schupp I.          BMC Genomics 8:399-399(2007) [PubMed: 17974005] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 3).  <u>Tissue:</u> Fetal brain.</p>
[8]	<p><b>"The DNA sequence, annotation and analysis of human chromosome 3."</b>          Muzny D.M., Scherer S.E., Kaul R., Wang J., Yu J., Sudbrak R., Buhay C.J., Chen R., Cree A., Ding Y., Dugan-Rocha S., Gill R., Gunaratne P., Harris R.A., Hawes A.C., Hernandez J., Hodgson A.V., Hume J.  Gibbs R.A.          Nature 440:1194-1198(2006) [PubMed: 16641997] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].</p>
[9]	<p>Mural R.J., Istrail S., Sutton G.G., Florea L., Halpern A.L., Mobarry C.M., Lippert R., Walenz B., Shatkay H., Dew I., Miller J.R., Flanigan M.J., Edwards N.J., Bolanos R., Fasulo D.,</p>

	<p>Halldorsson B.V., Hannenhalli S., Turner R.  Venter J.C.  Submitted (SEP-2005) to the EMBL/GenBank/DDBJ databases  Cited for: NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].</p>
[10]	<p><b>"The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC)."</b>  The MGC Project Team  Genome Res. 14:2121-2127(2004) [PubMed: 15489334] [Abstract]  Cited for: NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 2 AND 3).  Tissue: Fetal brain.</p>
[11]	<p><b>"Interaction properties of human mannan-binding lectin (MBL)-associated serine proteases-1 and -2, MBL-associated protein 19, and MBL."</b>  Thielens N.M., Cseh S., Thiel S., Vorup-Jensen T., Rossi V., Jensenius J.C., Arlaud G.J.  J. Immunol. 166:5068-5077(2001) [PubMed: 11290788] [Abstract]  Cited for: PROTEIN SEQUENCE OF 20-29 AND 449-458, SIGNAL SEQUENCE CLEAVAGE SITE, CLEAVAGE AT ARG-448, GLYCOSYLATION, HOMODIMERIZATION, INTERACTION WITH MBL2.</p>
[12]	<p><b>"Human serum mannose-binding lectin (MBL)-associated serine protease-1 (MASP-1): determination of levels in body fluids and identification of two forms in serum."</b>  Terai I., Kobayashi K., Matsushita M., Fujita T.  Clin. Exp. Immunol. 110:317-323(1997) [PubMed: 9367419] [Abstract]  Cited for: TISSUE SPECIFICITY.</p>
[13]	<p><b>"A second serine protease associated with mannan-binding lectin that activates complement."</b>  Thiel S., Vorup-Jensen T., Stover C.M., Schwaebler W.J., Laursen S.B., Poulsen K., Willis A.C., Eggleton P., Hansen S., Holmskov U., Reid K.B.M., Jensenius J.C.  Nature 386:506-510(1997) [PubMed: 9087411] [Abstract]  Cited for: INTERACTION WITH MBL2.  Tissue: Liver.</p>
[14]	<p><b>"Complement-activating complex of ficolin and mannose-binding lectin-associated serine protease."</b>  Matsushita M., Endo Y., Fujita T.  J. Immunol. 164:2281-2284(2000) [PubMed: 10679061] [Abstract]  Cited for: INTERACTION WITH FCN2.</p>
[15]	<p><b>"Interaction of C1q and mannan-binding lectin (MBL) with C1r, C1s, MBL-associated serine proteases 1 and 2, and the MBL-associated protein MAP19."</b>  Thiel S., Petersen S.V., Vorup-Jensen T., Matsushita M., Fujita T., Stover C.M., Schwaebler W.J., Jensenius J.C.  J. Immunol. 165:878-887(2000) [PubMed: 10878362] [Abstract]  Cited for: INTERACTION WITH MBL2.</p>
[16]	<p><b>"Proteolytic activities of two types of mannose-binding lectin-associated serine protease."</b>  Matsushita M., Thiel S., Jensenius J.C., Terai I., Fujita T.  J. Immunol. 165:2637-2642(2000) [PubMed: 10946292] [Abstract]  Cited for: CATALYTIC ACTIVITY, INTERACTION WITH SERPING1.</p>
[17]	<p><b>"Substrate specificities of recombinant mannan-binding lectin-associated serine proteases-1 and -2."</b>  Rossi V., Cseh S., Bally I., Thielens N.M., Jensenius J.C., Arlaud G.J.  J. Biol. Chem. 276:40880-40887(2001) [PubMed: 11527969] [Abstract]  Cited for: CATALYTIC ACTIVITY, BIOPHYSICOCHEMICAL PROPERTIES, ENZYME REGULATION.</p>
[18]	<p><b>"Activation of the lectin complement pathway by H-ficolin (Hakata antigen)."</b>  Matsushita M., Kuraya M., Hamasaki N., Tsujimura M., Shiraki H., Fujita T.  J. Immunol. 168:3502-3506(2002) [PubMed: 11907111] [Abstract]  Cited for: INTERACTION WITH FCN3.</p>

[19]	<p><b>"Characterization of the interaction between L-ficolin/p35 and mannan-binding lectin-associated serine proteases-1 and -2."</b>  Cseh S., Vera L., Matsushita M., Fujita T., Arlaud G.J., Thielens N.M.  J. Immunol. 169:5735-5743(2002) [PubMed: 12421953] [Abstract]  <u>Cited for:</u> INTERACTION WITH FCN2.</p>
[20]	<p><b>"Natural substrates and inhibitors of mannan-binding lectin-associated serine protease -1 and -2: a study on recombinant catalytic fragments."</b>  Ambrus G., Gal P., Kojima M., Szilagyi K., Balczer J., Antal J., Graf L., Laich A., Moffatt B.E., Schwaeble W., Sim R.B., Zavodszky P.  J. Immunol. 170:1374-1382(2003) [PubMed: 12538697] [Abstract]  <u>Cited for:</u> ENZYME REGULATION, BIOPHYSICOCHEMICAL PROPERTIES.</p>
[21]	<p><b>"Characterization of recombinant mannan-binding lectin-associated serine protease (MASP)-3 suggests an activation mechanism different from that of MASP-1 and MASP-2."</b>  Zundel S., Cseh S., Lacroix M., Dahl M.R., Matsushita M., Andrieu J.-P., Schwaeble W.J., Jensenius J.C., Fujita T., Arlaud G.J., Thielens N.M.  J. Immunol. 172:4342-4350(2004) [PubMed: 15034049] [Abstract]  <u>Cited for:</u> CHARACTERIZATION (ISOFORM 2), MUTAGENESIS OF SER-646, AUTOCATALYTIC CLEAVAGE.</p>
[22]	<p><b>"Human plasma N-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry."</b>  Liu T., Qian W.-J., Gritsenko M.A., Camp D.G. II, Monroe M.E., Moore R.J., Smith R.D.  J. Proteome Res. 4:2070-2080(2005) [PubMed: 16335952] [Abstract]  <u>Cited for:</u> GLYCOSYLATION [LARGE SCALE ANALYSIS] AT ASN-49; ASN-178; ASN-385; ASN-407 (ISOFORM 1), GLYCOSYLATION [LARGE SCALE ANALYSIS] AT ASN-533 AND ASN-599 (ISOFORM 2), MASS SPECTROMETRY.  <u>Tissue:</u> Plasma.</p>
[23]	<p><b>"Mannan-binding lectin-associated serine protease 3 cleaves synthetic peptides and insulin-like growth factor-binding protein 5."</b>  Cortesio C.L., Jiang W.  Arch. Biochem. Biophys. 449:164-170(2006) [PubMed: 16554018] [Abstract]  <u>Cited for:</u> CATALYTIC ACTIVITY (ISOFORM 2).</p>
[24]	<p><b>"Cooperation between MASP-1 and MASP-2 in the generation of C3 convertase through the MBL pathway."</b>  Moeller-Kristensen M., Thiel S., Sjoeholm A., Matsushita M., Jensenius J.C.  Int. Immunol. 19:141-149(2007) [PubMed: 17182967] [Abstract]  <u>Cited for:</u> FUNCTION (ISOFORMS 1 AND 2).</p>
[25]	<p><b>"Crystal structure of the CUB1-EGF-CUB2 domain of human MASP-1/3 and identification of its interaction sites with mannan-binding lectin and ficolins."</b>  Teillet F., Gaboriaud C., Lacroix M., Martin L., Arlaud G.J., Thielens N.M.  J. Biol. Chem. 283:25715-25724(2008) [PubMed: 18596036] [Abstract]  <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 20-295 IN COMPLEX WITH CALCIUM IONS, HOMODIMERIZATION, GLYCOSYLATION AT ASN-178, DISULFIDE BONDS, CALCIUM-BINDING SITES, INTERACTION WITH FCN2; FCN3 AND MBL2, MUTAGENESIS OF GLU-68; TYR-77; GLU-99; ASP-121; PHE-122; SER-123; GLU-125; HIS-237; GLU-239; TYR-244; GLU-262; SER-274; ASN-283 AND GLU-286.</p>
+	Additional computationally mapped references.



**Cross-references****Sequence databases**

<input checked="" type="radio"/> EMBL <input checked="" type="radio"/> GenBank <input checked="" type="radio"/> DDBJ	D17525 mRNA. Translation: BAA04477.1. D28593 mRNA. Translation: BAA05928.1. D61695 Genomic DNA. Translation: BAA34864.1. AB007617 Genomic DNA. Translation: BAA89206.1. AF284421 mRNA. Translation: AAK84071.1. AK291157 mRNA. Translation: BAF83846.1. AK291686 mRNA. Translation: BAF84375.1. AK304334 mRNA. Translation: BAG65179.1. CR749615 mRNA. Translation: CAH18409.1. AC007920 Genomic DNA. No translation available. CH471052 Genomic DNA. Translation: EAW78153.1. BC039724 mRNA. Translation: AAH39724.1. Different initiation. BC106945 mRNA. Translation: AAI06946.1. BC106946 mRNA. Translation: AAI06947.1.
--	---

IPI	IPI00216882. IPI00290283. IPI00299307. IPI00924973.
-----	--

PIR	I54763.
-----	---------

RefSeq	NP_001027019.1. NM_001031849.2. NP_001870.3. NM_001879.5. NP_624302.1. NM_139125.3.
--------	---

UniGene	Hs.89983.
---------	-----------

**3D structure databases**

<input checked="" type="radio"/> PDBe	Entry	Method	Resolution	Chain	Positions	PDBsum
<input checked="" type="radio"/> RCSB PDB			(Å)			
<input checked="" type="radio"/> PDBj	3DEM	X-ray	2.30	A/B	20-297	[»]
	3GOV	X-ray	2.55	A	298-448	[»]
				B	449-699	[»]

ProteinModelPortal	P48740.
--------------------	---------

SMR	P48740. Positions 26-444, 449-699.
-----	------------------------------------

ModBase	Search...
---------	-----------

**Protein-protein interaction databases**

STRING	P48740.
--------	---------

**Protein family/group databases**

MEROPS	S01.198.
--------	----------


**Proteomic databases**

PRIDE	P48740.
-------	---------

**Genome annotation databases**

Ensembl	ENST00000337774; ENSP00000336792; ENSG00000127241.
---------	--

GeneID	5648.
KEGG	hsa:5648.
UCSC	uc003frh.1. human.
<b>Organism-specific databases</b>	
CTD	5648.
GeneCards	GC03M184338.
HGNC	HGNC:6901. MASP1.
HPA	HPA001617. HPA009641.
MIM	600521. gene.
neXtProt	NX_P48740.
PharmGKB	PA30644.
GenAtlas	Search...
<b>Phylogenomic databases</b>	
eggNOG	prNOG16590.
GeneTree	ENSGT00560000076882.
HOVERGEN	HBG000559.
OMA	GPAPHML.
OrthoDB	EOG4001HP.
<b>Enzyme and pathway databases</b>	
Reactome	REACT_6900. Signaling in Immune system.
<b>Gene expression databases</b>	
ArrayExpress	P48740.
Bgee	P48740.
CleanEx	HS_MASP1.
Genevestigator	P48740.
GermOnline	ENSG00000127241. Homo sapiens.
<b>Family and domain databases</b>	
InterPro	IPR016060. Complement_control_module. IPR000859. CUB. IPR013032. EGF-like_reg_CS. IPR001881. EGF_Ca-bd. IPR013091. EGF_Ca-bd_2. IPR018097. EGF_Ca-bd_CS. IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. IPR000436. Sushi_SCR_CCP. [Graphical view]
Gene3D	G3DSA:2.10.70.10. Complement_control_module. 2 hits. G3DSA:2.60.120.290. CUB. 2 hits.

Pfam	PF00431. CUB. 2 hits. PF07645. EGF_CA. 1 hit. PF00084. Sushi. 2 hits. PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00032. CCP. 2 hits. SM00042. CUB. 2 hits. SM00179. EGF_CA. 1 hit. SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF57535. Complement_control_module. 2 hits. SSF49854. CUB. 2 hits. SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS00010. ASX_HYDROXYL. 1 hit. PS01180. CUB. 2 hits. PS00022. EGF_1. False negative. PS01186. EGF_2. 1 hit. PS50026. EGF_3. False negative. PS01187. EGF_CA. 1 hit. PS50923. SUSHI. 2 hits. PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
NextBio	21938.
SOURCE	Search...
<b>Entry information</b>	
Entry name	MASP1_HUMAN
Accession	Primary (citable) accession number: <b>P48740</b> Secondary accession number(s): A8K542  Q9UF09
Entry history	Integrated into February 1, 1996 UniProtKB/Swiss-Prot: Last sequence update: December 16, 2008 Last modified: April 5, 2011 This is version 116 of the entry and version 3 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
Disclaimer	Any medical or genetic information present in this entry is provided for research, educational and informational purposes only. It is not in any way intended to be used as a substitute for professional medical advice, diagnosis, treatment or care.
<b>Relevant documents</b>	
Human chromosome 3	

Human chromosome 3: entries, gene names and cross-references to MIM
Human entries with polymorphisms or disease mutations List of human entries with polymorphisms or disease mutations
Human polymorphisms and disease mutations Index of human polymorphisms and disease mutations
MIM cross-references Online Mendelian Inheritance in Man (MIM) cross-references in UniProtKB/Swiss-Prot
PDB cross-references Index of Protein Data Bank (PDB) cross-references
Peptidase families Classification of peptidase families and list of entries
SIMILARITY comments Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact





**P98064 (MASP1\_MOUSE) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified March 8, 2011. Version 106.

**Names and origin**

## Protein names

*Recommended name:***Mannan-binding lectin serine protease 1**

EC=3.4.21.-

*Alternative name(s):*

Complement factor MASP-3

Complement-activating component of Ra-reactive factor

Mannose-binding lectin-associated serine protease 1

Short name=MASP-1

Mannose-binding protein-associated serine protease

Ra-reactive factor serine protease p100

Short name=RaRF

Serine protease 5

Cleaved into the following 2 chains:

1. **Mannan-binding lectin serine protease 1 heavy chain**
2. **Mannan-binding lectin serine protease 1 light chain**

## Gene names

Name: **Masp1**

Synonyms: Crarf, Masp3

## Organism

**Mus musculus (Mouse)**

## Taxonomic identifier

10090 [NCBI]

## Taxonomic lineage

Eukaryota › Metazoa › Chordata › Craniata › Vertebrata ›  
Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires  
› Rodentia › Sciurognathi › Muroidea › Muridae › Murinae › Mus  
› Mus

**Protein attributes**

## Sequence length

704 AA.

## Sequence status

Complete.

## Sequence processing

The displayed sequence is further processed into a mature form.

## Protein existence

Evidence at protein level.

**General annotation (Comments)**

## Function

Functions in the lectin pathway of complement, which performs a key role in innate immunity by recognizing pathogens through patterns of sugar moieties and neutralizing them. The lectin pathway is triggered upon binding of mannan-binding lectin (MBL) and ficolins to sugar moieties which leads to activation of the associated proteases MASP1 and MASP2. Functions as an endopeptidase and may activate MASP2 or C2 or directly activate C3 the key component of complement reaction.

	Isoform 2 may have an inhibitory effect on the activation of the lectin pathway of complement or may cleave IGFBP5. <a href="#">Ref.7</a>
Enzyme regulation	Inhibited by SERPING1 and A2M <a href="#">By similarity</a> .
Subunit structure	Homodimer. Interacts with the oligomeric lectins MBL1, MBL2, FCN2 and FCN3; triggers the lectin pathway of complement through activation of C3. Interacts with SERPING1 <a href="#">By similarity</a> .
Subcellular location	Secreted <a href="#">Ref.1</a> .
Tissue specificity	Protein of the plasma which is primarily expressed by liver. <a href="#">Ref.7</a> <a href="#">Ref.1</a>
Post-translational modification	<p>The iron and 2-oxoglutarate dependent 3-hydroxylation of aspartate and asparagine is (R) stereospecific within EGF domains <a href="#">By similarity</a>.</p> <p>N-glycosylated. Some N-linked glycan are of the complex-type <a href="#">By similarity</a>.</p> <p>Autoproteolytic processing of the proenzyme produces the active enzyme composed on the heavy and the light chain held together by a disulfide bond. Isoform 1 but not isoform 2 is activated through autoproteolytic processing <a href="#">By similarity</a>.</p>
Disruption phenotype	Mice are smaller and more vulnerable indicating developmental and growth defects. Mice serum has low C4 and C3 cleavage activity together with low MASP2 activation. <a href="#">Ref.7</a>
Sequence similarities	<p>Belongs to the peptidase S1 family.</p> <p>Contains 2 CUB domains.</p> <p>Contains 1 EGF-like domain.</p> <p>Contains 1 peptidase S1 domain.</p> <p>Contains 2 Sushi (CCP/SCR) domains.</p>

## Ontologies

### Keywords

Biological process	Complement activation lectin pathway Immunity Innate immunity
Cellular component	Secreted
Coding sequence diversity	Alternative splicing
Domain	EGF-like domain Repeat Signal Sushi
Ligand	Calcium Metal-binding
Molecular function	Hydrolase Protease Serine protease
PTM	Autocatalytic cleavage Disulfide bond

Technical term	Glycoprotein Hydroxylation
	Direct protein sequencing
<b>Gene Ontology (GO)</b>	
Biological process	complement activation, lectin pathway Inferred from mutant phenotype <a href="#">Ref.7</a> . Source: UniProtKB
	proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	extracellular space Inferred from direct assay <a href="#">Ref.7</a> . Source: UniProtKB
Molecular function	calcium ion binding Inferred from sequence or structural similarity. Source: UniProtKB
	protein homodimerization activity Inferred from sequence or structural similarity. Source: UniProtKB
	serine-type endopeptidase activity Inferred from direct assay <a href="#">Ref.7</a> . Source: UniProtKB
Complete GO annotation...	

Alternative products

This entry describes 2 isoforms produced by **alternative splicing**. [\[Align\]](#) [\[Select\]](#)

Isoform 1 (identifier: **P98064-1**)

Also known as: MASP-1;

*This isoform has been chosen as the 'canonical' sequence. All positional information in this entry refers to it. This is also the sequence that appears in the downloadable versions of the entry.*



Isoform 2 (identifier: **P98064-2**)




Also known as: MASP-3;

*The sequence of this isoform differs from the canonical sequence as follows:*  
443-443: V → QPSRALPNLV...RAVRDLQVER  
444-704: Missing.











Note: Glycosylated on Asn-538 and Asn-604 (By similarity).

Sequence annotation (Features)
















Feature key	Position(s)	Length	Description	Graphical view
<b>Molecule processing</b>				
 Signal peptide	1 – 24	24		

	Chain	25 – 704	680	Mannan-binding lectin serine protease 1		
	Chain	25 – 453	429	Mannan-binding lectin serine protease 1 heavy chain		
	Chain	454 – 704	251	Mannan-binding lectin serine protease 1 light chain		

**Regions**

	Domain	25 – 143	119	CUB 1		
	Domain	144 – 187	44	EGF-like; calcium-binding <small>By similarity</small>		
	Domain	190 – 302	113	CUB 2		
	Domain	304 – 369	66	Sushi 1		
	Domain	370 – 439	70	Sushi 2		
	Domain	454 – 701	248	Peptidase S1		
	Region	25 – 305	281	Interaction with MBL1 <small>By similarity</small>		
	Region	25 – 283	259	Interaction with FCN2 <small>By similarity</small>		
	Region	25 – 189	165	Homodimerization <small>By similarity</small>		
	Region	25 – 189	165	Interaction with MBL2 <small>By similarity</small>		

**Sites**

	Active site	495	1	Charge relay system <small>By similarity</small>		
	Active site	557	1	Charge relay system <small>By similarity</small>		
	Active site	651	1	Charge relay system <small>By similarity</small>		
	Metal binding	73	1	Calcium 1 <small>By similarity</small>		
	Metal binding	81	1	Calcium 1 <small>By similarity</small>		
	Metal binding	126	1	Calcium 1 <small>By similarity</small>		
	Metal binding	128	1	Calcium 1; via carbonyl oxygen <small>By similarity</small>		
	Metal binding	144	1	Calcium 2 <small>By similarity</small>		
	Metal binding	145	1	Calcium 2; via carbonyl oxygen <small>By similarity</small>		
	Metal binding	147	1	Calcium 2 <small>By similarity</small>		
	Metal binding	164	1	Calcium 2 <small>By similarity</small>		
	Metal binding	165	1	Calcium 2; via carbonyl oxygen <small>By similarity</small>		
	Metal binding	168	1	Calcium 2; via carbonyl oxygen <small>By similarity</small>		
	Metal binding	240	1	Calcium 3 <small>By similarity</small>		
	Metal binding	250	1	Calcium 3 <small>By similarity</small>		



<input type="checkbox"/>	Metal binding	287	1	Calcium 3 <span>By similarity</span>	
<input type="checkbox"/>	Metal binding	289	1	Calcium 3; via carbonyl oxygen <span>By similarity</span>	
<input type="checkbox"/>	Site	453 – 454	2	Cleavage; by autolysis <span>By similarity</span>	

**Amino acid modifications**

<input type="checkbox"/>	Modified residue	164	1	(3R)-3-hydroxyasparagine <span>Potential</span>	
<input type="checkbox"/>	Glycosylation	54	1	N-linked (GlcNAc...) <span>Potential</span>	
<input type="checkbox"/>	Glycosylation	183	1	N-linked (GlcNAc...) <span>Potential</span>	
<input type="checkbox"/>	Glycosylation	390	1	N-linked (GlcNAc...) <span>Potential</span>	
<input type="checkbox"/>	Glycosylation	412	1	N-linked (GlcNAc...) <span>Potential</span>	
<input type="checkbox"/>	Disulfide bond	78 ↔ 96		<span>By similarity</span>	
<input type="checkbox"/>	Disulfide bond	148 ↔ 162		<span>By similarity</span>	
<input type="checkbox"/>	Disulfide bond	158 ↔ 171		<span>By similarity</span>	
<input type="checkbox"/>	Disulfide bond	173 ↔ 186		<span>By similarity</span>	
<input type="checkbox"/>	Disulfide bond	190 ↔ 217		<span>By similarity</span>	
<input type="checkbox"/>	Disulfide bond	247 ↔ 265		<span>By similarity</span>	
<input type="checkbox"/>	Disulfide bond	306 ↔ 354		<span>By similarity</span>	
<input type="checkbox"/>	Disulfide bond	334 ↔ 367		<span>By similarity</span>	
<input type="checkbox"/>	Disulfide bond	372 ↔ 419		<span>By similarity</span>	
<input type="checkbox"/>	Disulfide bond	402 ↔ 437		<span>By similarity</span>	
<input type="checkbox"/>	Disulfide bond	441 ↔ 577		Interchain (between heavy and light chains) <span>Potential</span>	
<input type="checkbox"/>	Disulfide bond	480 ↔ 496		<span>By similarity</span>	
<input type="checkbox"/>	Disulfide bond	619 ↔ 636		<span>By similarity</span>	
<input type="checkbox"/>	Disulfide bond	647 ↔ 677		<span>By similarity</span>	

**Natural variations**

<input type="checkbox"/>	Alternative sequence	443	1	V → QPSRALPNLVKRIIGGRNAE LGLFPWQALIVVEDTSRVPN DKWFGSGALLSESWILTAH VLRSQRRDNTVIPVSKEHVT VYLGLHDVRDKSGAVNSSAA RVILHPDFNIQNYNHDIALV QLQKPVPLGAHVMPICLPRP EPEGPAPHMLGLVAGWGISN PNVTVDEILSGTRTLSDVL QYVKLPVVS HAECKASYESR SGNYSVTENMFCAGYYEGGK DTCLGDSSGGAFFVIFDEMSQH WWAQGLVSWGGPEECGSKQV	
--------------------------	----------------------	-----	---	--	--

<input type="checkbox"/>	Alternative sequence	444 – 704	261	YGVYTKVSNYVDWLWEEMNS PRAVRDLQVER in isoform 2.
<input type="checkbox"/>	Alternative sequence	444 – 704	261	Missing in isoform 2.

**Experimental info**

<input type="checkbox"/>	Sequence conflict	257	1	G → A in AAI31638. <a href="#">Ref.5</a>
<input type="checkbox"/>	Sequence conflict	345	1	E → G in BAA03944. <a href="#">Ref.1</a>
<input type="checkbox"/>	Sequence conflict	345	1	E → G in BAB69688. <a href="#">Ref.2</a>
<input type="checkbox"/>	Sequence conflict	428	1	E → K in BAA03944. <a href="#">Ref.1</a>
<input type="checkbox"/>	Sequence conflict	465	1	M → T in AAN39850. <a href="#">Ref.2</a>

**Sequences**

Sequence	Length	Mass (Da)
<input type="checkbox"/> <b>Isoform 1 (MASP-1) [UniParc].</b>	FASTA	704

79,968

Last modified April 14, 2009. Version 2.  
Checksum: 4CF0B17916C10961

```

      10      20      30      40      50      60
MRFLSFWRLL LYHALCLALP EVSAHTVELN EMFGQIQSPG YPDSYPSDSE VTWNITVPEG

      70      80      90     100     110     120
FRIKLYFMHF NLESSYLCEY DYVKVETEDQ VLATFCGRET TDTEQTPGQE VVLSPGTFMS

     130     140     150     160     170     180
VTFRSDFSNE ERFTGFDAHY MAVDVDECKE REDEELSCDH YCHNYIGGYI CSCRFGYILH

     190     200     210     220     230     240
TDNRTCVRVE SGNLFTQRTG TITSPDYPNP YPKSSECSYT IDLEEGFMVS LQFEDIFDIE

     250     260     270     280     290     300
DHPEVPCPYD YIKIKAGSKV WGPFCGEKSP EPISTQTHSV QILFRSDNSG ENRGWRLSYR

     310     320     330     340     350     360
AAGNECPKLQ PPVYGKIEPS QAVYSFKDQV LVSCDTGYKV LKDNEVMDTF QIECLKDGAW

     370     380     390     400     410     420
SNKIPTCKIV DCGAPAGLKH GLVTFSTRNN LTTYKSEIRY SCQQPYKML HNTTGVYTCS

     430     440     450     460     470     480
AHGTWTNEVL KRSLPTCLPV CGVPKFSRKQ ISRIFNGRPA QKGTMPWIAM LSHLNGQPFC

     490     500     510     520     530     540
GGSLLGSNWV LTAACHLHQS LDPEEPTLHS SYLLSPSDFK IIMGKHWRRR SDEDEQHLHV

     550     560     570     580     590     600
KRTTLHPLYN PSTFENDLGL VELSESPRLN DFVMPVCLPE QPSTEGTMVI VSGWGKQFLQ

     610     620     630     640     650     660
RFPENLMEIE IPIVNSDTCQ EAYTPLKKKV TKDMICAGEK EGGKDACAGD SGGPMVTKDA

     670     680     690     700
ERDQWYLVGV VSWGEDCGKK DRYGVYSIY PNKDWIQRIT GVRN
```

« Hide

<input type="checkbox"/> <b>Isoform 2 (MASP-3).</b>	FASTA	733	82,446
Checksum: 6B3B54D5D3F5822B			
Show »			

## References

	« Hide 'large scale' references
[1]	<p><b>"A 100-kDa protein in the C4-activating component of Ra-reactive factor is a new serine protease having module organization similar to C1r and C1s."</b>            Takayama Y., Takada F., Takahashi A., Kawakami M.            J. Immunol. 152:2308-2316(1994) [PubMed: 8133044] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), PARTIAL PROTEIN SEQUENCE, SUBCELLULAR LOCATION, TISSUE SPECIFICITY.  <u>Strain:</u> BALB/c.  <u>Tissue:</u> Liver.</p>
[2]	<p><b>"Murine serine proteases MASP-1 and MASP-3, components of the lectin pathway activation complex of complement, are encoded by a single structural gene."</b>            Stover C.M., Lynch N.J., Dahl M.R., Hanson S., Takahashi M., Frankenberger M., Ziegler-Heitbrock L., Eperon I., Thiel S., Schwaebler W.J.            Genes Immun. 4:374-384(2003) [PubMed: 12847554] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA / MRNA] (ISOFORM 2), ALTERNATIVE SPLICING (ISOFORM 1).  <u>Strain:</u> BALB/c.  <u>Tissue:</u> Liver.</p>
[3]	<p><b>"The transcriptional landscape of the mammalian genome."</b>            Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N., Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K., Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M., Davis M.J., Hayashizaki Y.            Science 309:1559-1563(2005) [PubMed: 16141072] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).  <u>Strain:</u> C57BL/6J.  <u>Tissue:</u> Testis.</p>
[4]	<p>Mural R.J., Adams M.D., Myers E.W., Smith H.O., Venter J.C.            Submitted (JUL-2005) to the EMBL/GenBank/DDBJ databases  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].</p>
[5]	<p><b>"The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC)."</b>            The MGC Project Team            Genome Res. 14:2121-2127(2004) [PubMed: 15489334] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).</p>
[6]	<p><b>"Presence of a serine protease in the complement-activating component of the complement-dependent bactericidal factor, RaRF, in mouse serum."</b>            Takahashi A., Takayama Y., Hatsuse H., Kawakami M.            Biochem. Biophys. Res. Commun. 190:681-687(1993) [PubMed: 8439319] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] OF 465-704 (ISOFORM 1), PARTIAL PROTEIN SEQUENCE.  <u>Strain:</u> BALB/c.  <u>Tissue:</u> Liver.</p>
[7]	<p><b>"Mannose-binding lectin (MBL)-associated serine protease (MASP)-1 contributes to activation of the lectin complement pathway."</b>            Takahashi M., Iwaki D., Kanno K., Ishida Y., Xiong J., Matsushita M., Endo Y., Miura S., Ishii N., Sugamura K., Fujita T.            J. Immunol. 180:6132-6138(2008) [PubMed: 18424734] [Abstract]  <u>Cited for:</u> FUNCTION, DISRUPTION PHENOTYPE, TISSUE SPECIFICITY.</p>
+	Additional computationally mapped references.

**Cross-references****Sequence databases**

<ul style="list-style-type: none"> <li>⊙ EMBL</li> <li>⊙ GenBank</li> <li>⊙ DDBJ</li> </ul>	D16492 mRNA. Translation: BAA03944.1. AB049755 mRNA. Translation: BAB69688.1. AY135527, AY135525 Genomic DNA. Translation: AAN39850.1. AK031598 mRNA. Translation: BAC27469.1. CH466521 Genomic DNA. Translation: EDK97670.1. CH466521 Genomic DNA. Translation: EDK97671.1. BC131637 mRNA. Translation: AAI31638.1. BC131638 mRNA. Translation: AAI31639.1.
---	---

IPI	IPI00475209. IPI00848780.
PIR	PC1235.
RefSeq	NP_032581.2. NM_008555.2.
UniGene	Mm.1213.

**3D structure databases**

ProteinModelPortal	P98064.
SMR	P98064. Positions 32-449, 454-704.
ModBase	Search...

**Protein-protein interaction databases**

STRING	P98064.
--------	---------

**Protein family/group databases**

MEROPS	S01.198.
--------	----------

**Proteomic databases**

PRIDE	P98064.
-------	---------

**Genome annotation databases**

Ensembl	ENSMUST00000023605; ENSMUSP00000023605; ENSMUSG00000022887.
GeneID	17174.
KEGG	mmu:17174.
UCSC	uc007ytr.1. mouse.

**Organism-specific databases**

CTD	17174.
MGI	MGI:88492. Masp1.

**Phylogenomic databases**

eggNOG	roNOG05755.
--------	-------------




GeneTree	ENSGT00560000076882.
HOVERGEN	HBG000559.
OMA	GPAPHML.
<b>Gene expression databases</b>	
ArrayExpress	P98064.
Bgee	P98064.
CleanEx	MM_MASP1.
Genevestigator	P98064.
GermOnline	ENSMUSG00000022887. Mus musculus.
<b>Family and domain databases</b>	
InterPro	IPR016060. Complement_control_module. IPR000859. CUB. IPR013032. EGF-like_reg_CS. IPR001881. EGF_Ca-bd. IPR013091. EGF_Ca-bd_2. IPR018097. EGF_Ca-bd_CS. IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. IPR000436. Sushi_SCR_CCP. [Graphical view]
Gene3D	G3DSA:2.10.70.10. Complement_control_module. 2 hits. G3DSA:2.60.120.290. CUB. 2 hits.
Pfam	PF00431. CUB. 2 hits. PF07645. EGF_CA. 1 hit. PF00084. Sushi. 2 hits. PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00032. CCP. 2 hits. SM00042. CUB. 2 hits. SM00179. EGF_CA. 1 hit. SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF57535. Complement_control_module. 2 hits. SSF49854. CUB. 2 hits. SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS00010. ASX_HYDROXYL. 1 hit. PS01180. CUB. 2 hits. PS00022. EGF_1. False negative. PS01186. EGF_2. 1 hit. PS50026. EGF_3. False negative. PS01187. EGF_CA. 1 hit. PS50923. SUSHI. 2 hits. PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...

**Other Resources**

SOURCE

Search...

**Entry information**

Entry name	MASP1_MOUSE
Accession	Primary (citable) accession number: <b>P98064</b> Secondary accession number(s): A2RRH8  Q920S0
Entry history	Integrated into February 1, 1996 UniProtKB/Swiss-Prot: Last sequence update: April 14, 2009 Last modified: March 8, 2011 This is version 106 of the entry and version 2 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program

**Relevant documents**

MGD cross-references

Mouse Genome Database (MGD) cross-references in UniProtKB/Swiss-Prot

Peptidase families

Classification of peptidase families and list of entries

SIMILARITY comments

Index of protein domains and families

© 2002–2011 UniProt Consortium | License &amp; Disclaimer | Contact

EMBL-EBI



**P48740 (MASP1\_HUMAN)** ☆ Reviewed, UniProtKB/Swiss-Prot

Last modified April 5, 2011. Version 116.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Mannan-binding lectin serine protease 1</b> EC=3.4.21.- <i>Alternative name(s):</i> Complement factor MASP-3 Complement-activating component of Ra-reactive factor Mannose-binding lectin-associated serine protease 1 Short name=MASP-1 Mannose-binding protein-associated serine protease Ra-reactive factor serine protease p100 Short name=RaRF Serine protease 5 <u>Cleaved into the following 2 chains:</u> 1. <b>Mannan-binding lectin serine protease 1 heavy chain</b> 2. <b>Mannan-binding lectin serine protease 1 light chain</b>
Gene names	Name: <b>MASP1</b> Synonyms:CRARF, CRARF1, PRSS5
Organism	<b>Homo sapiens (Human)</b> [Complete proteome]
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo

**Protein attributes**

Sequence length	699 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Functions in the lectin pathway of complement, which performs a key role in innate immunity by recognizing pathogens through patterns of sugar moieties and neutralizing them. The lectin pathway is triggered upon binding of mannan-binding lectin (MBL) and ficolins to sugar moieties which leads to activation of the associated proteases MASP1 and MASP2. Functions as an endopeptidase and may activate MASP2 or C2 or directly activate C3 the key component of complement reaction. Isoform 2 may have an inhibitory effect on the activation of the
----------	--

	lectin pathway of complement or may cleave IGFBP5. <a href="#">Ref.5</a> <a href="#">Ref.24</a>
Enzyme regulation	Inhibited by SERPING1 and A2M. <a href="#">Ref.17</a> <a href="#">Ref.20</a>
Subunit structure	Homodimer. Interacts with the oligomeric lectins MBL2, FCN2 and FCN3; triggers the lectin pathway of complement through activation of C3. Interacts with SERPING1. <a href="#">Ref.11</a> <a href="#">Ref.13</a> <a href="#">Ref.14</a> <a href="#">Ref.15</a> <a href="#">Ref.16</a> <a href="#">Ref.18</a> <a href="#">Ref.19</a> <a href="#">Ref.25</a>
Subcellular location	Secreted <a href="#">Ref.5</a>
Tissue specificity	Protein of the plasma which is primarily expressed by liver. <a href="#">Ref.5</a> <a href="#">Ref.1</a> <a href="#">Ref.2</a> <a href="#">Ref.12</a>
Post-translational modification	The iron and 2-oxoglutarate dependent 3-hydroxylation of aspartate and asparagine is (R) stereospecific within EGF domains <a href="#">By similarity</a> .  N-glycosylated. Some N-linked glycan are of the complex-type <a href="#">By similarity</a> . <a href="#">Ref.11</a> <a href="#">Ref.25</a> <a href="#">Ref.22</a>  Autoproteolytic processing of the proenzyme produces the active enzyme composed on the heavy and the light chain held together by a disulfide bond. Isoform 1 but not isoform 2 is activated through autoproteolytic processing.
Sequence similarities	Belongs to the peptidase S1 family. Contains 2 CUB domains. Contains 1 EGF-like domain. Contains 1 peptidase S1 domain. Contains 2 Sushi (CCP/SCR) domains.
Biophysicochemical properties	<u>Kinetic parameters:</u>  $K_M=0.10$ mM for Ac-Gly-Lys-OMe (at 30 degrees Celsius) <a href="#">Ref.17</a> <a href="#">Ref.20</a>  $K_M=310$ $\mu$ M for Bz-Arg-OEt (at 30 degrees Celsius)  $K_M=4.8$ $\mu$ M for C2 (at 37 degrees Celsius)
Sequence caution	The sequence AAH39724.1 differs from that shown. Reason: Erroneous initiation.

## Ontologies

### Keywords

Biological process	Complement activation lectin pathway Immunity Innate immunity
Cellular component	Secreted
Coding sequence diversity	Alternative splicing Polymorphism
Domain	EGF-like domain Repeat Signal Sushi



Ligand	Calcium Metal-binding
Molecular function	Hydrolase Protease Serine protease
PTM	Autocatalytic cleavage Disulfide bond Glycoprotein Hydroxylation
Technical term	3D-structure Complete proteome Direct protein sequencing
<b>Gene Ontology (GO)</b>	
Biological process	complement activation, lectin pathway Inferred from mutant phenotype <a href="#">Ref.24</a> . Source: UniProtKB  negative regulation of complement activation Inferred from direct assay <a href="#">Ref.5</a> . Source: UniProtKB  proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	extracellular space Inferred from direct assay <a href="#">Ref.5</a> . Source: UniProtKB
Molecular function	calcium ion binding Inferred from direct assay <a href="#">Ref.25</a> . Source: UniProtKB  calcium-dependent protein binding Inferred from physical interaction <a href="#">Ref.19</a> <a href="#">Ref.19</a> <a href="#">Ref.25</a> . Source: UniProtKB  protein binding Inferred from physical interaction <a href="#">Ref.16</a> <a href="#">Ref.5</a> <a href="#">Ref.18</a> . Source: UniProtKB  protein homodimerization activity Inferred from physical interaction <a href="#">Ref.25</a> . Source: UniProtKB  serine-type endopeptidase activity Inferred from direct assay <a href="#">Ref.17</a> . Source: UniProtKB
Complete GO annotation...	
<b>Alternative products</b>	
This entry describes 4 isoforms produced by <b>alternative splicing</b> . <a href="#">[Align]</a> <a href="#">[Select]</a>	
<b>Isoform 1 (identifier: P48740-1)</b>	
<i>This isoform has been chosen as the 'canonical' sequence. All positional information in this entry refers to it. This is also the sequence that appears in the downloadable versions of the entry.</i>	
<b>Isoform 2 (identifier: P48740-2)</b>	
Also known as: MASP-3;	

The sequence of this isoform differs from the canonical sequence as follows:

435-435: V → ECGQPSRSLP...QSVVEPQVER

436-699: Missing.

Note: Glycosylated on Asn-533 and Asn-599.

#### Isoform 3 (identifier: P48740-3)

The sequence of this isoform differs from the canonical sequence as follows:

364-380: IVDCRAPGELEHGLITF → KNEIDLESELKSEQVTE

381-699: Missing.

#### Isoform 4 (identifier: P48740-4)

















The sequence of this isoform differs from the canonical sequence as follows:

1-113: Missing.

435-435: V → ECGQPSRSLP...QSVVEPQVER

436-699: Missing.

### Sequence annotation (Features)

	Feature key	Position(s)	Length	Description	Graphical view
<b>Molecule processing</b>					
	Signal peptide	1 – 19	19	<a href="#">Ref.11</a>	
	Chain	20 – 699	680	Mannan-binding lectin serine protease 1	
	Chain	20 – 448	429	Mannan-binding lectin serine protease 1 heavy chain	
	Chain	449 – 699	251	Mannan-binding lectin serine protease 1 light chain	
<b>Regions</b>					
	Domain	20 – 138	119	CUB 1	
	Domain	139 – 182	44	EGF-like; calcium-binding	
	Domain	185 – 297	113	CUB 2	
	Domain	299 – 364	66	Sushi 1	
	Domain	365 – 434	70	Sushi 2	
	Domain	449 – 696	248	Peptidase S1	
	Region	20 – 278	259	Interaction with FCN2	
	Region	20 – 184	165	Homodimerization <a href="#">By similarity</a>	
	Region	20 – 184	165	Interaction with MBL2	
<b>Sites</b>					
	Active site	490	1	Charge relay system <a href="#">By similarity</a>	
	Active site	552	1	Charge relay system <a href="#">By similarity</a>	
	Active site	646	1	Charge relay system <a href="#">By similarity</a>	

<input type="checkbox"/>	Metal binding	68	1	Calcium 1	
<input type="checkbox"/>	Metal binding	76	1	Calcium 1	
<input type="checkbox"/>	Metal binding	121	1	Calcium 1	
<input type="checkbox"/>	Metal binding	123	1	Calcium 1; via carbonyl oxygen	
<input type="checkbox"/>	Metal binding	139	1	Calcium 2	
<input type="checkbox"/>	Metal binding	140	1	Calcium 2; via carbonyl oxygen	
<input type="checkbox"/>	Metal binding	142	1	Calcium 2	
<input type="checkbox"/>	Metal binding	159	1	Calcium 2	
<input type="checkbox"/>	Metal binding	160	1	Calcium 2; via carbonyl oxygen	
<input type="checkbox"/>	Metal binding	163	1	Calcium 2; via carbonyl oxygen	
<input type="checkbox"/>	Metal binding	235	1	Calcium 3	
<input type="checkbox"/>	Metal binding	245	1	Calcium 3	
<input type="checkbox"/>	Metal binding	282	1	Calcium 3	
<input type="checkbox"/>	Metal binding	284	1	Calcium 3; via carbonyl oxygen	
<input type="checkbox"/>	Site	448 – 449	2	Cleavage; by autolysis	
<b>Amino acid modifications</b>					
<input type="checkbox"/>	Modified residue	159	1	(3R)-3-hydroxyasparagine Potential	
<input type="checkbox"/>	Glycosylation	49	1	N-linked (GlcNAc...) Ref.22	
<input type="checkbox"/>	Glycosylation	178	1	N-linked (GlcNAc...) Ref.25 Ref.22	
<input type="checkbox"/>	Glycosylation	385	1	N-linked (GlcNAc...) Ref.22	
<input type="checkbox"/>	Glycosylation	407	1	N-linked (GlcNAc...) Ref.22	
<input type="checkbox"/>	Disulfide bond	73 ↔ 91		Ref.25	
<input type="checkbox"/>	Disulfide bond	143 ↔ 157		Ref.25	
<input type="checkbox"/>	Disulfide bond	153 ↔ 166		Ref.25	
<input type="checkbox"/>	Disulfide bond	168 ↔ 181		Ref.25	
<input type="checkbox"/>	Disulfide bond	185 ↔ 212		Ref.25	
<input type="checkbox"/>	Disulfide bond	242 ↔ 260		Ref.25	
<input type="checkbox"/>	Disulfide bond	301 ↔ 349		By similarity	
<input type="checkbox"/>	Disulfide bond	329 ↔ 362		By similarity	
<input type="checkbox"/>	Disulfide bond	367 ↔ 414		By similarity	
<input type="checkbox"/>	Disulfide bond	397 ↔ 432		By similarity	
<input type="checkbox"/>	Disulfide bond	436 ↔ 572		Interchain (between heavy and light chains) Potential	
<input type="checkbox"/>	Disulfide bond	475 ↔ 491		By similarity	
<input type="checkbox"/>	Disulfide bond	614 ↔ 631		By similarity	
<input type="checkbox"/>	Disulfide bond	642 ↔ 672		By similarity	

Natural variations					
<input checked="" type="checkbox"/>	Alternative sequence	1 – 113	113	Missing in isoform 4.	I
<input checked="" type="checkbox"/>	Alternative sequence	364 – 380	17	IVDCR...GLITF → KNEIDLESELKSEQVTE in isoform 3.	II
<input checked="" type="checkbox"/>	Alternative sequence	381 – 699	319	Missing in isoform 3.	I
<input type="checkbox"/>	Alternative sequence	435	1	V → ECGQPSRSLPSLVKRIIGGR NAEPGLFPWQALIVEDTSR VPNDKWFGSGALLSASWILT AAHVLRSQRRDTTVIPVSKE HVTVYLGLHDVRDKSGAVNS SAARVVLHPDFNIQNYNHDI ALVQLQEPVPLGPHVMPVCL PRLEPEGPAPHMLGLVAGWG ISNPNTVDEIISGTRTLS DVLQYVKLPVVPFAECKTSY ESRSGNYSVTENMFCAGYYE GGKDTCLGDSGGAFVIFDDL SQRWVVQGLVSWGGPEECGS KQVYGVYTKVSNYVDWVWEQ MGLPQSVVEPQVER in isoform 2 and isoform 4.	I
<input checked="" type="checkbox"/>	Alternative sequence	436 – 699	264	Missing in isoform 2 and isoform 4.	I
<input type="checkbox"/>	Natural variant	21	1	T → I. [dbSNP:rs1062049]	I
<input type="checkbox"/>	Natural variant	568	1	V → A. [dbSNP:rs13322090]	
<input type="checkbox"/>	Natural variant	679	1	G → R. [dbSNP:rs3774266]	
Experimental info					
<input type="checkbox"/>	Mutagenesis	68	1	E → A or Q: Partial loss of interaction with FCN2, FCN3 and MBL2. [Ref.25]	I
<input type="checkbox"/>	Mutagenesis	77	1	Y → A: Partial loss of interaction with FCN2, FCN3 and MBL2. [Ref.25]	I
<input type="checkbox"/>	Mutagenesis	99	1	E → A: Partial loss of interaction with FCN2, FCN3 and MBL2. [Ref.25]	I
<input type="checkbox"/>	Mutagenesis	121	1	D → A or N: Loss of interaction with FNC2 and FCN3 and partial loss of interaction with MBL2. [Ref.25]	I
<input type="checkbox"/>	Mutagenesis	122	1	F → A: Partial loss of interaction with FCN2, FCN3 and MBL2. [Ref.25]	I
<input type="checkbox"/>	Mutagenesis	123	1	S → A: Partial loss of interaction with FCN2, FCN3 and MBL2. [Ref.25]	I



<input type="checkbox"/>	Mutagenesis	125	1	E → A: Partial loss of interaction with FCN2, FCN3 and MBL2. <a href="#">Ref.25</a>	
<input type="checkbox"/>	Mutagenesis	237	1	H → A: Loss of interaction with FCN2, FCN3 and MBL2. <a href="#">Ref.25</a>	
<input type="checkbox"/>	Mutagenesis	239	1	E → A: Partial loss of interaction with FCN2, FCN3 and MBL2. <a href="#">Ref.25</a>	
<input type="checkbox"/>	Mutagenesis	244	1	Y → A: Loss of interaction with FCN2, FCN3 and MBL2. <a href="#">Ref.25</a>	
<input type="checkbox"/>	Mutagenesis	262	1	E → A: Partial loss of interaction with FCN2, FCN3 and MBL2. <a href="#">Ref.25</a>	
<input type="checkbox"/>	Mutagenesis	274	1	S → A: Partial loss of interaction with FCN2 and FCN3. No effect on interaction with MBL2. <a href="#">Ref.25</a>	
<input type="checkbox"/>	Mutagenesis	283	1	N → A: Partial loss of interaction with FCN2, FCN3 and MBL2. <a href="#">Ref.25</a>	
<input type="checkbox"/>	Mutagenesis	286	1	E → A: Partial loss of interaction with FCN2, FCN3 and MBL2. <a href="#">Ref.25</a>	
<input type="checkbox"/>	Mutagenesis	646	1	S → A: No autoproteolytic processing. <a href="#">Ref.21</a>	
<input type="checkbox"/>	Sequence conflict	2	1	R → K in BAF84375. <a href="#">Ref.6</a>	
<input type="checkbox"/>	Sequence conflict	89	1	T → A in CAH18409. <a href="#">Ref.7</a>	
<input type="checkbox"/>	Sequence conflict	232	1	F → L in BAF84375. <a href="#">Ref.6</a>	
<input type="checkbox"/>	Sequence conflict	235	1	E → Q in BAA05928. <a href="#">Ref.2</a>	
<input type="checkbox"/>	Sequence conflict	235	1	E → Q in BAA34864. <a href="#">Ref.3</a>	
<input type="checkbox"/>	Sequence conflict	285	1	G → A in BAA05928. <a href="#">Ref.2</a>	
<input type="checkbox"/>	Sequence conflict	285	1	G → A in BAA34864. <a href="#">Ref.3</a>	
<input type="checkbox"/>	Sequence conflict	285	1	G → A in BAA89206. <a href="#">Ref.4</a>	
<input type="checkbox"/>	Sequence conflict	392	1	E → G in BAF83846. <a href="#">Ref.6</a>	
<input type="checkbox"/>	Sequence conflict	499	1	E → G in BAA05928. <a href="#">Ref.2</a>	
<input type="checkbox"/>	Sequence conflict	499	1	E → K in BAA04477. <a href="#">Ref.1</a>	
<input type="checkbox"/>	Sequence conflict	499	1	E → K in BAA89206. <a href="#">Ref.4</a>	
<input type="checkbox"/>	Sequence conflict	527	1	D → A in BAA34864. <a href="#">Ref.3</a>	
<input type="checkbox"/>	Sequence conflict	543	1	Q → K in BAA04477. <a href="#">Ref.1</a>	
<input type="checkbox"/>	Sequence conflict	552	1	D → V in BAA34864. <a href="#">Ref.3</a>	
<input type="checkbox"/>	Sequence conflict	643	1	A → S in BAA04477. <a href="#">Ref.1</a>	

**Secondary structure**

1. ....

■ Helix ■ Strand ■ Turn

Details...


Sequences

Sequence	Length	Mass (Da)
<div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div>Isoform 1 [UniParc].</div><div>Last modified December 16, 2008. Version 3.</div><div>Checksum: 5B37C7FB9F51FD1D</div></div></div><div><div>FASTA</div><div>699</div><div>79,247</div></div></div></div>	699	79,247
<div><div><div>MRWLLLYAL CFSLSKASAH TVELNNMFGQ IQSPGYPDSY PSDSEVTWNI TVPDGFRIKL</div><div>YFMHFNLESS YLCEYDYVKV ETEDQVLATF CGRETTDTEQ TPGQEVVLSP GSFMSITFRS</div><div>DFSNEERFTG FDAHYMAVDV DECKEREDEE LSCDHYCHNY IGGYYCSCRf GYILHTDNRT</div><div>CRVECSDNLF TQRTGVITSP DFPNPYPKSS ECLYTIELEE GFMVNLQFED IFDIEDHPEV</div><div>PCPYDIYIKI VGPKVLGPFC GEKAPEPIST QSHSVLILFH SDNSGENRGW RLSYRAAGNE</div><div>CPELQPPVHG KIEPSQAKYF FKDQVLVSCD TGYKVLKDNV EMDTFQIECL KDGTWSNKIP</div><div>TCKIVDCRAP GELEHGLITF STRNNLTYYK SEIKYSCQEP YYKMLNNNTG IYTCSAQGVW</div><div>MNKVLGRSLP TCLPVCGLPK FSRKLMARIF NGRPAQKGTt PWIAMLSHLN GQPF CGGSLl</div><div>GSSWIVTAAH CLHQSLDPED PTLRSDLLS PSDFKIILGK HWRLRSDENE QHLGVKHTTL</div><div>HPQYDPNTFE NDVALVELLE SPVLNAFVMP ICLPEGPQQE GAMVIVSGWG KQFLQRFPEt</div><div>LMEIEIPIVD HSTCQKAYAP LKKKVTRDMI CAGEKEGGKD ACAGDSGGPM VTLNRERGQW</div><div>YLVGTVSWGd DCGKKDRYGV YSYIHHNKDW IQRVTGVRN</div></div></div>		
« Hide		
<div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div>Isoform 2 (MASP-3).</div><div>Checksum: 09B5297A6C14283A</div><div>Show »</div></div></div><div><div>FASTA</div><div>728</div><div>81,860</div></div></div></div>	728	81,860
<div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div>Isoform 3.</div><div>Checksum: DDED114311A62714</div><div>Show »</div></div></div><div><div>FASTA</div><div>380</div><div>43,640</div></div></div></div>	380	43,640
<div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div>Isoform 4.</div><div>Checksum: EEE63886709340FA</div><div>Show »</div></div></div><div><div>FASTA</div><div>615</div><div>68,918</div></div></div></div>	615	68,918

References

« Hide 'large scale' references



[1]	<p><b>"A new member of the C1s family of complement proteins found in a bactericidal factor, Ra-reactive factor, in human serum."</b>          Takada F., Takayama Y., Hatsuse H., Kawakami M.          Biochem. Biophys. Res. Commun. 196:1003-1009(1993) [PubMed: 8240317] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), TISSUE SPECIFICITY.  <u>Tissue:</u> Liver.</p>
[2]	<p><b>"Molecular characterization of a novel serine protease involved in activation of the complement system by mannose-binding protein."</b>          Sato T., Endo Y., Matsushita M., Fujita T.          Int. Immunol. 6:665-669(1994) [PubMed: 8018603] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), TISSUE SPECIFICITY.  <u>Tissue:</u> Fetal liver.</p>
[3]	<p><b>"Exon structure of the gene encoding the human mannose-binding protein-associated serine protease light chain: comparison with complement C1r and C1s genes."</b>          Endo Y., Sato T., Matsushita M., Fujita T.          Int. Immunol. 8:1355-1358(1996) [PubMed: 8921412] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA].  <u>Tissue:</u> Placenta.</p>
[4]	<p><b>"Gene structure of the P100 serine-protease component of the human Ra-reactive factor."</b>          Takayama Y., Takada F., Nowatari M., Kawakami M., Matsu-ura N.          Mol. Immunol. 36:505-514(1999) [PubMed: 10475605] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA].</p>
[5]	<p><b>"MASP-3 and its association with distinct complexes of the mannan-binding lectin complement activation pathway."</b>          Dahl M.R., Thiel S., Matsushita M., Fujita T., Willis A.C., Christensen T., Vorup-Jensen T., Jensenius J.C.          Immunity 15:127-135(2001) [PubMed: 11485744] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 2), PROTEIN SEQUENCE OF 450-474; 506-526; 539-555; 577-590; 613-621 AND 679-695 (ISOFORM 2), FUNCTION, SUBCELLULAR LOCATION, TISSUE SPECIFICITY.  <u>Tissue:</u> Liver.</p>
[6]	<p><b>"Complete sequencing and characterization of 21,243 full-length human cDNAs."</b>          Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., Sugano S.          Nat. Genet. 36:40-45(2004) [PubMed: 14702039] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 2 AND 4).  <u>Tissue:</u> Placenta, Teratocarcinoma and Trachea.</p>
[7]	<p><b>"The full-ORF clone resource of the German cDNA consortium."</b>          Bechtel S., Rosenfelder H., Duda A., Schmidt C.P., Ernst U., Wellenreuther R., Mehrle A., Schuster C., Bahr A., Bloeker H., Heubner D., Hoerlein A., Michel G., Wedler H., Koehrer K., Ottenwaelder B., Poustka A., Wiemann S., Schupp I.          BMC Genomics 8:399-399(2007) [PubMed: 17974005] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 3).  <u>Tissue:</u> Fetal brain.</p>
[8]	<p><b>"The DNA sequence, annotation and analysis of human chromosome 3."</b>          Muzny D.M., Scherer S.E., Kaul R., Wang J., Yu J., Sudbrak R., Buhay C.J., Chen R., Cree A., Ding Y., Dugan-Rocha S., Gill R., Gunaratne P., Harris R.A., Hawes A.C., Hernandez J., Hodgson A.V., Hume J., Gibbs R.A.          Nature 440:1194-1198(2006) [PubMed: 16641997] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].</p>
[9]	<p>Mural R.J., Istrail S., Sutton G.G., Florea L., Halpern A.L., Mobarry C.M., Lippert R., Walenz B., Shatkay H., Dew I., Miller J.R., Flanigan M.J., Edwards N.J., Bolanos R., Fasulo D.,</p>

	<p>Halldorsson B.V., Hannenhalli S., Turner R.  Venter J.C.  Submitted (SEP-2005) to the EMBL/GenBank/DDBJ databases  Cited for: NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].</p>
[10]	<p><b>"The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC)."</b>  The MGC Project Team  Genome Res. 14:2121-2127(2004) [PubMed: 15489334] [Abstract]  Cited for: NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 2 AND 3).  Tissue: Fetal brain.</p>
[11]	<p><b>"Interaction properties of human mannan-binding lectin (MBL)-associated serine proteases-1 and -2, MBL-associated protein 19, and MBL."</b>  Thielens N.M., Cseh S., Thiel S., Vorup-Jensen T., Rossi V., Jensenius J.C., Arlaud G.J.  J. Immunol. 166:5068-5077(2001) [PubMed: 11290788] [Abstract]  Cited for: PROTEIN SEQUENCE OF 20-29 AND 449-458, SIGNAL SEQUENCE CLEAVAGE SITE, CLEAVAGE AT ARG-448, GLYCOSYLATION, HOMODIMERIZATION, INTERACTION WITH MBL2.</p>
[12]	<p><b>"Human serum mannose-binding lectin (MBL)-associated serine protease-1 (MASP-1): determination of levels in body fluids and identification of two forms in serum."</b>  Terai I., Kobayashi K., Matsushita M., Fujita T.  Clin. Exp. Immunol. 110:317-323(1997) [PubMed: 9367419] [Abstract]  Cited for: TISSUE SPECIFICITY.</p>
[13]	<p><b>"A second serine protease associated with mannan-binding lectin that activates complement."</b>  Thiel S., Vorup-Jensen T., Stover C.M., Schwaebler W.J., Laursen S.B., Poulsen K., Willis A.C., Eggleton P., Hansen S., Holmskov U., Reid K.B.M., Jensenius J.C.  Nature 386:506-510(1997) [PubMed: 9087411] [Abstract]  Cited for: INTERACTION WITH MBL2.  Tissue: Liver.</p>
[14]	<p><b>"Complement-activating complex of ficolin and mannose-binding lectin-associated serine protease."</b>  Matsushita M., Endo Y., Fujita T.  J. Immunol. 164:2281-2284(2000) [PubMed: 10679061] [Abstract]  Cited for: INTERACTION WITH FCN2.</p>
[15]	<p><b>"Interaction of C1q and mannan-binding lectin (MBL) with C1r, C1s, MBL-associated serine proteases 1 and 2, and the MBL-associated protein MASP-19."</b>  Thiel S., Petersen S.V., Vorup-Jensen T., Matsushita M., Fujita T., Stover C.M., Schwaebler W.J., Jensenius J.C.  J. Immunol. 165:878-887(2000) [PubMed: 10878362] [Abstract]  Cited for: INTERACTION WITH MBL2.</p>
[16]	<p><b>"Proteolytic activities of two types of mannose-binding lectin-associated serine protease."</b>  Matsushita M., Thiel S., Jensenius J.C., Terai I., Fujita T.  J. Immunol. 165:2637-2642(2000) [PubMed: 10946292] [Abstract]  Cited for: CATALYTIC ACTIVITY, INTERACTION WITH SERPING1.</p>
[17]	<p><b>"Substrate specificities of recombinant mannan-binding lectin-associated serine proteases-1 and -2."</b>  Rossi V., Cseh S., Bally I., Thielens N.M., Jensenius J.C., Arlaud G.J.  J. Biol. Chem. 276:40880-40887(2001) [PubMed: 11527969] [Abstract]  Cited for: CATALYTIC ACTIVITY, BIOPHYSICOCHEMICAL PROPERTIES, ENZYME REGULATION.</p>
[18]	<p><b>"Activation of the lectin complement pathway by H-ficolin (Hakata antigen)."</b>  Matsushita M., Kuraya M., Hamasaki N., Tsujimura M., Shiraki H., Fujita T.  J. Immunol. 168:3502-3506(2002) [PubMed: 11907111] [Abstract]  Cited for: INTERACTION WITH FCN3.</p>



[19]	<p><b>"Characterization of the interaction between L-ficolin/p35 and mannan-binding lectin-associated serine proteases-1 and -2."</b>  Cseh S., Vera L., Matsushita M., Fujita T., Arlaud G.J., Thielens N.M.  J. Immunol. 169:5735-5743(2002) [PubMed: 12421953] [Abstract]  Cited for: INTERACTION WITH FCN2.</p>
[20]	<p><b>"Natural substrates and inhibitors of mannan-binding lectin-associated serine protease -1 and -2: a study on recombinant catalytic fragments."</b>  Ambrus G., Gal P., Kojima M., Szilagyi K., Balczer J., Antal J., Graf L., Laich A., Moffatt B.E., Schwaeble W., Sim R.B., Zavodszky P.  J. Immunol. 170:1374-1382(2003) [PubMed: 12538697] [Abstract]  Cited for: ENZYME REGULATION, BIOPHYSICOCHEMICAL PROPERTIES.</p>
[21]	<p><b>"Characterization of recombinant mannan-binding lectin-associated serine protease (MASP)-3 suggests an activation mechanism different from that of MASP-1 and MASP-2."</b>  Zundel S., Cseh S., Lacroix M., Dahl M.R., Matsushita M., Andrieu J.-P., Schwaeble W.J., Jensenius J.C., Fujita T., Arlaud G.J., Thielens N.M.  J. Immunol. 172:4342-4350(2004) [PubMed: 15034049] [Abstract]  Cited for: CHARACTERIZATION (ISOFORM 2), MUTAGENESIS OF SER-646, AUTOCATALYTIC CLEAVAGE.</p>
[22]	<p><b>"Human plasma N-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry."</b>  Liu T., Qian W.-J., Gritsenko M.A., Camp D.G. II, Monroe M.E., Moore R.J., Smith R.D.  J. Proteome Res. 4:2070-2080(2005) [PubMed: 16335952] [Abstract]  Cited for: GLYCOSYLATION [LARGE SCALE ANALYSIS] AT ASN-49; ASN-178; ASN-385; ASN-407 (ISOFORM 1), GLYCOSYLATION [LARGE SCALE ANALYSIS] AT ASN-533 AND ASN-599 (ISOFORM 2), MASS SPECTROMETRY.  Tissue: Plasma.</p>
[23]	<p><b>"Mannan-binding lectin-associated serine protease 3 cleaves synthetic peptides and insulin-like growth factor-binding protein 5."</b>  Cortesio C.L., Jiang W.  Arch. Biochem. Biophys. 449:164-170(2006) [PubMed: 16554018] [Abstract]  Cited for: CATALYTIC ACTIVITY (ISOFORM 2).</p>
[24]	<p><b>"Cooperation between MASP-1 and MASP-2 in the generation of C3 convertase through the MBL pathway."</b>  Moeller-Kristensen M., Thiel S., Sjoeholm A., Matsushita M., Jensenius J.C.  Int. Immunol. 19:141-149(2007) [PubMed: 17182967] [Abstract]  Cited for: FUNCTION (ISOFORMS 1 AND 2).</p>
[25]	<p><b>"Crystal structure of the CUB1-EGF-CUB2 domain of human MASP-1/3 and identification of its interaction sites with mannan-binding lectin and ficolins."</b>  Teillet F., Gaboriaud C., Lacroix M., Martin L., Arlaud G.J., Thielens N.M.  J. Biol. Chem. 283:25715-25724(2008) [PubMed: 18596036] [Abstract]  Cited for: X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 20-295 IN COMPLEX WITH CALCIUM IONS, HOMODIMERIZATION, GLYCOSYLATION AT ASN-178, DISULFIDE BONDS, CALCIUM-BINDING SITES, INTERACTION WITH FCN2; FCN3 AND MBL2, MUTAGENESIS OF GLU-68; TYR-77; GLU-99; ASP-121; PHE-122; SER-123; GLU-125; HIS-237; GLU-239; TYR-244; GLU-262; SER-274; ASN-283 AND GLU-286.</p>
+	Additional computationally mapped references.

**Cross-references****Sequence databases**

 EMBL  
 GenBank  
 DDBJ

D17525 mRNA. Translation: BAA04477.1.  
 D28593 mRNA. Translation: BAA05928.1.  
 D61695 Genomic DNA. Translation: BAA34864.1.  
 AB007617 Genomic DNA. Translation: BAA89206.1.  
 AF284421 mRNA. Translation: AAK84071.1.  
 AK291157 mRNA. Translation: BAF83846.1.  
 AK291686 mRNA. Translation: BAF84375.1.  
 AK304334 mRNA. Translation: BAG65179.1.  
 CR749615 mRNA. Translation: CAH18409.1.  
 AC007920 Genomic DNA. No translation available.  
 CH471052 Genomic DNA. Translation: EAW78153.1.  
 BC039724 mRNA. Translation: AAH39724.1. Different initiation.  
 BC106945 mRNA. Translation: AAI06946.1.  
 BC106946 mRNA. Translation: AAI06947.1.

IPI  
 IPI00216882.  
 IPI00290283.  
 IPI00299307.  
 IPI00924973.

PIR I54763.

RefSeq  
 NP\_001027019.1. NM\_001031849.2.  
 NP\_001870.3. NM\_001879.5.  
 NP\_624302.1. NM\_139125.3.

UniGene Hs.89983.

**3D structure databases**

 PDBe  
 RCSB PDB  
 PDBj

Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
3DEM	X-ray	2.30	A/B	20-297	[»]
3GOV	X-ray	2.55	A	298-448	[»]
			B	449-699	[»]

ProteinModelPortal P48740.

SMR P48740. Positions 26-444, 449-699.

ModBase Search...

**Protein-protein interaction databases**

STRING P48740.

**Protein family/group databases**

MEROPS S01.198.


**Proteomic databases**

PRIDE P48740.

**Genome annotation databases**

Ensembl ENST00000337774; ENSP00000336792; ENSG00000127241.

GeneID	5648.
KEGG	hsa:5648.
UCSC	uc003frh.1. human.
<b>Organism-specific databases</b>	
CTD	5648.
GeneCards	GC03M184338.
HGNC	HGNC:6901. MASP1.
HPA	HPA001617. HPA009641.
MIM	600521. gene.
neXtProt	NX_P48740.
PharmGKB	PA30644.
GenAtlas	Search...
<b>Phylogenomic databases</b>	
eggNOG	prNOG16590.
GeneTree	ENSGT00560000076882.
HOVERGEN	HBG000559.
OMA	GPAPHML.
OrthoDB	EOG4001HP.
<b>Enzyme and pathway databases</b>	
Reactome	REACT_6900. Signaling in Immune system.
<b>Gene expression databases</b>	
ArrayExpress	P48740.
Bgee	P48740.
CleanEx	HS_MASP1.
Genevestigator	P48740.
GermOnline	ENSG00000127241. Homo sapiens.
<b>Family and domain databases</b>	
InterPro	IPR016060. Complement_control_module. IPR000859. CUB. IPR013032. EGF-like_reg_CS. IPR001881. EGF_Ca-bd. IPR013091. EGF_Ca-bd_2. IPR018097. EGF_Ca-bd_CS. IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. IPR000436. Sushi_SCR_CCP. [Graphical view]
Gene3D	G3DSA:2.10.70.10. Complement_control_module. 2 hits. G3DSA:2.60.120.290. CUB. 2 hits.

Pfam	PF00431. CUB. 2 hits. PF07645. EGF_CA. 1 hit. PF00084. Sushi. 2 hits. PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN.
SMART	SM00032. CCP. 2 hits. SM00042. CUB. 2 hits. SM00179. EGF_CA. 1 hit. SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF57535. Complement_control_module. 2 hits. SSF49854. CUB. 2 hits. SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS00010. ASX_HYDROXYL. 1 hit. PS01180. CUB. 2 hits. PS00022. EGF_1. False negative. PS01186. EGF_2. 1 hit. PS50026. EGF_3. False negative. PS01187. EGF_CA. 1 hit. PS50923. SUSHI. 2 hits. PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
NextBio	21938.
SOURCE	Search...
<b>Entry information</b>	
Entry name	MASP1_HUMAN
Accession	Primary (citable) accession number: <b>P48740</b> Secondary accession number(s): A8K542  Q9UF09
Entry history	Integrated into February 1, 1996 UniProtKB/Swiss-Prot: Last sequence update: December 16, 2008 Last modified: April 5, 2011 This is version 116 of the entry and version 3 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
Disclaimer	Any medical or genetic information present in this entry is provided for research, educational and informational purposes only. It is not in any way intended to be used as a substitute for professional medical advice, diagnosis, treatment or care.
<b>Relevant documents</b>	
Human chromosome 3	



**Q96P40 (Q96P40\_HUMAN)** ☆ Unreviewed, UniProtKB/TrEMBL

Last modified March 2, 2010. Version 19.

**Names and origin**

Protein names	Submitted name: ADAMTS13 <a href="#">EMBL AAL11094.1</a>
Gene names	Name:ADAMTS13 <a href="#">EMBL AAL11094.1</a>
Organism	Homo sapiens (Human) <a href="#">EMBL AAL11094.1</a>
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo

**Protein attributes**

Sequence length	45 AA.
Sequence status	Fragment.
Protein existence	Evidence at transcript level.

**Sequence annotation (Features)**

Feature key	Position (s)	Length	Description	Graphical view	Feature identifier
-------------	-----------------	--------	-------------	----------------	-----------------------

**Experimental info**

<input type="checkbox"/>	Non-terminal residue	1	1	<a href="#">EMBL AAL11094.1</a>	
<input type="checkbox"/>	Non-terminal residue	45	1	<a href="#">EMBL AAL11094.1</a>	

**Sequences**

Sequence	Length	Mass (Da)
<input type="checkbox"/> Q96P40 [UniParc]. Last modified December 1, 2001. Version 1. Checksum: 2D4AA2E2F39EC633 <div><div>10203040</div><div>LFVCRFTGGM ARSMATSPAQ TSPSPTSSL S HGRPGCGPLC VGPAR</div></div> « Hide	FASTA 45	4,535

**References**

- [1] "Mutations in a member of the ADAMTS gene family cause thrombotic thrombocytopenic purpura."

Levy G.G., Nichols W.C., Lian E.C., Foroud T., McClintick J.N., McGee B.M., Yang A.Y., Siemieniak D.R., Stark K.R., Gruppo R., Sarode R., Shurin S.B., Chandrasekaran V., Stabler S.P., Sabio H., Bouhassira E.E., Upshaw J.D. Jr., Ginsburg D., Tsai H.-M.  
Nature 413:488-494(2001) [PubMed: 11586351] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE.  
Tissue: Liver [\[EMBL AAL11094.1\]](#).

## Cross-references

### Sequence databases

-  EMBL
-  GenBank
-  DDBJ

AF414400 mRNA. Translation: AAL11094.1.

### 3D structure databases

ModBase

Search...

### Family and domain databases

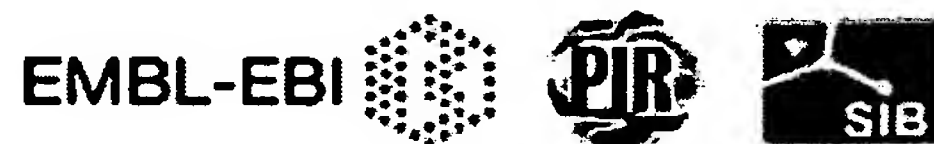
ProtoNet

Search...

## Entry information

Entry name	Q96P40_HUMAN
Accession	Primary (citable) accession number: <b>Q96P40</b>
Entry history	<p>Integrated into      December 1, 2001 UniProtKB/TrEMBL: Last sequence      December 1, 2001 update: Last modified:      March 2, 2010 This is version 19 of the entry and version 1 of the sequence. [Complete history]</p>
Entry status	Unreviewed (UniProtKB/TrEMBL)
Disclaimer	Any medical or genetic information present in this entry is provided for research, educational and informational purposes only. It is not in any way intended to be used as a substitute for professional medical advice, diagnosis, treatment or care.

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



## Protein

Translations of Life

Display Settings: GenPept

## von Willebrand factor-cleaving protease [Homo sapiens]

GenBank: BAB69487.2

[FASTA](#) [Graphics](#)

## Go to:

LOCUS BAB69487 1427 aa linear PRI 17-OCT-2001  
DEFINITION von Willebrand factor-cleaving protease [Homo sapiens].  
ACCESSION BAB69487  
VERSION BAB69487.2 GI:16117338  
DBSOURCE accession AB069698.2  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Soejima,K., Mimura,N., Hirashima,M., Maeda,H., Hamamoto,T.,  
Nakagaki,T. and Nozaki,C.  
TITLE A novel human metalloprotease synthesized in the liver and secreted  
into the blood: possibly, the von Willebrand factor-cleaving  
protease?  
JOURNAL J. Biochem. 130 (4), 475-480 (2001)  
PUBMED 11574066  
REFERENCE 2 (residues 1 to 1427)  
AUTHORS Soejima,K. and Mimura,N.  
TITLE Direct Submission  
JOURNAL Submitted (08-AUG-2001) Kenji Soejima, The Chemo-Sero-Therapeutic  
Research Institute, First Research Department; Kyokushi, Kikuchi,  
Kumamoto 869-1298, Japan (E-mail:soejima@kaketsuken.or.jp,  
Tel:81-968-37-3100(ex.5250), Fax:81-968-37-3616)  
COMMENT On Oct 12, 2001 this sequence version replaced gi:15990913.  
FEATURES  
source 1..1427  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="9"  
/map="9q34"  
/tissue\_type="liver"  
Protein 1..1427  
/product="von Willebrand factor-cleaving protease"  
Region 80..283  
/region\_name="ZnMc\_ADAMTS\_like"  
/note="Zinc-dependent metalloprotease, ADAMTS\_like  
subgroup. ADAMs (A Disintegrin And Metalloprotease) are  
glycoproteins, which play roles in cell signaling, cell  
fusion, and cell-cell interactions. This particular  
subfamily represents domain architectures...; cd04273"  
/db\_xref="CDD:58574"  
Site order(224..225,228,234)  
/site\_type="active"  
/db\_xref="CDD:58574"  
Region 387..439  
/region\_name="TSP\_1"  
/note="Thrombospondin type 1 domain; cl02514"  
/db\_xref="CDD:154951"  
CDS 1..1427  
/gene="vWF-CP"  
/coded\_by="AB069698.2:445..4728"  
/note="ADAMTS13;  
a disintegrin-like and metalloprotease (reprolysin type)  
with thrombospondin type 1 motif, 13;  
vWFCP"  
ORIGIN  
1 mhqhrhprarc pplcvagila cgflilgcwgp shfqqsclqa lepqavssyl spgaplkgrp  
61 pspgfgqrgrq rqrreaaggil hlelllvavgp dvfqahqedt eryvltlnli gaellrdpsl  
121 gaqfrvhlvk mviltepega pnitanltss llsvcgwsqt inpedtdtpg hadlvlyitr  
181 fdlelpgdgnr qvrgvtqlgg acsptwscli tedtgfdlgr tiaheighsf glehdgapgs  
241 gcgpgshvma sdgaapragl awspcsrrql lsllsagrar cvwdpprpqp gsaghppdaq  
301 pglyysaneq crvafgpkav actfarehld mcqalschtd pldqsscsrl lvplldgtec  
361 gvekwcskgr crslveltpi aavhgrwssw gprspcsrsc gggvvtrrrq cnpnrpafgg  
421 racvgadlga emcntqacek tqlefmssqc artdgqplrs spggasfyhw gaavphsqqd  
481 alcrhmcrai gesfimkrqd sfldgtrcmp sgpredgtls lcvsgscrtf gcdgrmdsqq  
541 vwdrcqvcgg dnstcsprkg sftagrarey vtflvtvtnl tsvyianhrp lfthlavrig  
601 gryvvagkms ispnttypsl ledgrveyrv altdedrlprl eeiriwglq edadiqvyr  
661 ygeeygnltr pditftyfqp kprqawvwaa vrgpcsvscg aglrwnvysc ldqarkelve  
721 tvqcqgsqgp pawpeacvle pcppywavgd fgpcsvscgg glrerpvrvc eaqgsllktl  
781 pparcragaq qpavaletcn pqpccparvev sepssctsag gaglalenet cvpgadglea  
841 pvtgpggsvd eklpapepcv gmscppgwggh ldatsageka pspwgsirtg aqaahvwtpa  
901 agscsvscgr glmelrflcm dsalrvpvqe elcglaskpg srrevcqavp cparwqykla  
961 acsvscgrgv vrrilycara hgeddgeeil ldtqcqglpr pepqeacsle pcprrwkvms  
1021 lgpcsvscgl gtarrsvacv qldqgqdvvev deaacaalvr peasvpclia dcytrwhvgt  
1081 wmcsvscgd giqrirdtcl gpqacqavpva dfcqhlpkp tvrgcwagpc vgggtpslvp  
1141 heeaaapgrt tatpagasle wsgargllfs papqprllp gpqensvqss acgrqhlept  
1201 gtidmrpggq adcavaigrp lgevvtlrvl esslncsagd mlllwgrltw rkmcrrklldm  
1261 tfsstktntlv vrgrcgrpgg gvllrygsqql apettyfrecd mqlfgpwgei vpslspats  
1321 naggcrlfin vaphariaih alatinmgagt eganasayili rdthslrtta fhgqqvlywe  
1381 sessqaemef segflkaqas lrgqywtlqs wvpemqdpqs wkgkegt

//

**P08473 (NEP\_HUMAN) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 146.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Neprilysin</b> EC=3.4.24.11 <i>Alternative name(s):</i> Atriopeptidase Common acute lymphocytic leukemia antigen Short name=CALLA Enkephalinase Neutral endopeptidase 24.11 Short name=NEP Short name=Neutral endopeptidase Skin fibroblast elastase Short name=SFE CD_antigen=CD10
Gene names	Name: <b>MME</b> Synonyms: EPN
Organism	<b>Homo sapiens (Human)</b> [Complete proteome]
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo

**Protein attributes**

Sequence length	750 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Thermolysin-like specificity, but is almost confined on acting on polypeptides of up to 30 amino acids. Biologically important in the destruction of opioid peptides such as Met- and Leu-enkephalins by cleavage of a Gly-Phe bond. Able to cleave angiotensin-1, angiotensin-2 and angiotensin 1-9. Involved in the degradation of atrial natriuretic factor (ANF). Displays UV-inducible elastase activity toward skin preelastic and elastic fibers. <a href="#">Ref.9</a> <a href="#">Ref.12</a> <a href="#">Ref.14</a>
Catalytic activity	Preferential cleavage of polypeptides between hydrophobic residues, particularly with Phe or Tyr at P1'.
Cofactor	Binds 1 zinc ion per subunit. <a href="#">Ref.16</a> <a href="#">Ref.17</a>
Enzyme regulation	Inhibited in a dose dependent manner by opiorphin.



Subcellular location	Cell membrane; Single-pass type II membrane protein.
Miscellaneous	Important cell surface marker in the diagnostic of human acute lymphocytic leukemia.
Sequence similarities	Belongs to the peptidase M13 family.
Biophysicochemical properties	<u>Kinetic parameters:</u> $K_M=55.1 \mu\text{M}$ for angiotensin-1 <span style="border: 1px solid black; padding: 0 2px;">Ref.12</span> $K_M=179 \mu\text{M}$ for angiotensin-2 $K_M=111.4 \mu\text{M}$ for angiotensin 1-9
Sequence caution	The sequence CAA30157.1 differs from that shown. Reason: Erroneous initiation.

## Ontologies

### Keywords

Cellular component	Cell membrane Membrane
Domain	Signal-anchor Transmembrane Transmembrane helix
Ligand	Metal-binding Zinc
Molecular function	Hydrolase Metalloprotease Protease
PTM	Disulfide bond Glycoprotein
Technical term	3D-structure Complete proteome

### Gene Ontology (GO)

Biological process	cell-cell signaling Traceable author statement. Source: ProtInc proteolysis Traceable author statement. Source: ProtInc
Cellular component	integral to plasma membrane Traceable author statement. Source: ProtInc
Molecular function	metal ion binding Inferred from electronic annotation. Source: UniProtKB-KW metalloendopeptidase activity Inferred from electronic annotation. Source: InterPro protein binding Inferred from physical interaction. Source: IntAct

Complete GO annotation...

## Binary interactions

With	Entry	#Exp.	IntAct	Notes
HSPA1A	P08107	1	EBI-353759,EBI-629985	
HSPB1	P04792	1	EBI-353759,EBI-352682	
NDRG1	Q92597	1	EBI-353759,EBI-716486	

## Sequence annotation (Features)

	Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
<b>Molecule processing</b>						
<input type="checkbox"/>	Initiator methionine	1	1	Removed	I	
<input checked="" type="checkbox"/>	Chain	2 – 750	749	Neprilysin		PRO_0000078
<b>Regions</b>						
<input checked="" type="checkbox"/>	Topological domain	2 – 28	27	Cytoplasmic Potential	I	
<input checked="" type="checkbox"/>	Transmembrane	29 – 51	23	Helical; Signal- anchor for type II membrane protein; Potential	II	
<input checked="" type="checkbox"/>	Topological domain	52 – 750	699	Extracellular Potential	I	
<input type="checkbox"/>	Motif	16 – 23	8	Stop- transfer sequence Potential	I	
<b>Sites</b>						
<input type="checkbox"/>	Active site	591	1		I	
<input type="checkbox"/>	Active site	651	1	Proton donor Ref.10	I	
<input type="checkbox"/>	Metal binding	584	1	Zinc; catalytic	I	
<input type="checkbox"/>	Metal binding	588	1	Zinc; catalytic	I	
<input type="checkbox"/>	Metal binding	647	1	Zinc; catalytic	I	

<input type="checkbox"/>	Binding site	103	1	Substrate carboxyl By similarity			
<b>Amino acid modifications</b>							
<input type="checkbox"/>	Glycosylation	145	1	N-linked (GlcNAc...) Ref.16 Ref.17 Ref.11 Ref.15			
<input type="checkbox"/>	Glycosylation	285	1	N-linked (GlcNAc...) Ref.11			
<input type="checkbox"/>	Glycosylation	325	1	N-linked (GlcNAc...) Ref.16 Ref.17 Ref.15			
<input type="checkbox"/>	Glycosylation	628	1	N-linked (GlcNAc...) Ref.16 Ref.17 Ref.15			
<input type="checkbox"/>	Disulfide bond	57 ↔ 62		Ref.16 Ref.17			
<input type="checkbox"/>	Disulfide bond	80 ↔ 735		Ref.16 Ref.17			
<input type="checkbox"/>	Disulfide bond	88 ↔ 695		Ref.16 Ref.17			
<input type="checkbox"/>	Disulfide bond	143 ↔ 411		Ref.16 Ref.17			
<input type="checkbox"/>	Disulfide bond	234 ↔ 242		Ref.16 Ref.17			
<input type="checkbox"/>	Disulfide bond	621 ↔ 747		Ref.16 Ref.17			
<b>Experimental info</b>							
<input type="checkbox"/>	Sequence conflict	26	1	P → R in AAA51915. Ref.4			
<input type="checkbox"/>	Sequence conflict	44	1	T → R in AAA51915. Ref.4			
<input type="checkbox"/>	Sequence conflict	81	1	T → R in AAA51915. Ref.4			

<input type="checkbox"/>	Sequence conflict	304	1	T → R in AAA51915. Ref.4	1
--------------------------	-------------------	-----	---	--------------------------------	---

## Secondary structure

1 .

■ Helix ■ Strand ■ Turn

Details...

## Sequences

Sequence	Length	Mass (Da)
<input checked="" type="checkbox"/> P08473 [UniParc]. Last modified January 23, 2007. Version 2. Checksum: BCF3827C39898630	FASTA 750	85,514

```

      10      20      30      40      50      60
MGKSESQMDI TDINTPKPKK KQRWTPLEIS LSVLVLLLT IAVTMIALYA TYDDGICKSS

      70      80      90     100     110     120
DCIKSAARLI QNMDATTEPC TDFFKYACGG WLKRNVIPET SSRYGNFDIL RDELEVVLKD

     130     140     150     160     170     180
VLQEPKTEDI VAVQKAKALY RSCINESAID SRGGEPLLKL LPDIYGWPVA TENWEQKYGA

     190     200     210     220     230     240
SWTAEKAIAQ LNSKYGKKVL INLFVGTDDK NSVNHVIHID QPRLGLPSRD YYECTGIYKE

     250     260     270     280     290     300
ACTAYVDFMI SVARLIRQEE RLPIDENQLA LEMNKVMELE KEIANATAKP EDRNDPMLLY

     310     320     330     340     350     360
NKMTLAQIQN NFSLEINGK PFWLNFTNEI MSTVNISITN EEDVVVYAPE YLTKLKPILT

     370     380     390     400     410     420
KYSARDLQNL MSWRFIMDLV SSSLRTYKES RNAFRKALYG TTSETATWRR CANYVNGNME

     430     440     450     460     470     480
NAVGRLYVEA AFAGESKHVV EDLIAQIREV FIQTLDDLTV MDAETKKRAE EKALAIKERI

     490     500     510     520     530     540
GYPDDIVSND NKLNNEYLEL NYKEDEYFEN IIQNLKFSQS KQLKKLREKV DKDEWISGAA

     550     560     570     580     590     600
VVNAFYSSGR NQIVFPAGIL QPPFFSAQQS NSLNYGGIGM VIGHEITHGF DDNGRNFNKD

     610     620     630     640     650     660
GDLVDWWTQQ SASNFKEQSQ CMVYQYGNFS WDLAGGQHLN GINTLGENIA DNGGLGQAYR

     670     680     690     700     710     720
AYQNYIKKNG EEKLLPGLDL NHKQLFFLNF AQVWCITYRP EYAVNSIKTD VHSPGNFRII

     730     740     750
GTLQNSAEFS EAFHCRKNSY MNPEKKCRVW
```

« Hide

## References

« Hide 'large scale' references

- [1] "Common acute lymphocytic leukemia antigen is identical to neutral endopeptidase."  
Letarte M., Vera S., Tran R., Addis J.B.L., Onizuka R.J., Quackenbush E.J., Jongeneel C.V.,



	McInnes R.R. J. Exp. Med. 168:1247-1253(1988) [PubMed: 2971756] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA]. <u>Tissue:</u> Kidney.
[2]	<b>"Molecular cloning of the common acute lymphoblastic leukemia antigen (CALLA) identifies a type II integral membrane protein."</b> Shipp M.A., Richardson N.E., Sayre P.H., Brown N.R., Masteller E.L., Clayton L.K., Ritz J., Reinherz E.L. Proc. Natl. Acad. Sci. U.S.A. 85:4819-4823(1988) [PubMed: 2968607] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].
[3]	<b>"Organization of the gene encoding common acute lymphoblastic leukemia antigen (neutral endopeptidase 24.11): multiple minixons and separate 5' untranslated regions."</b> D'Adamio L., Shipp M.A., Masteller E.L., Reinherz E.L. Proc. Natl. Acad. Sci. U.S.A. 86:7103-7107(1989) [PubMed: 2528730] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA].
[4]	<b>"Complete sequencing and characterization of 21,243 full-length human cDNAs."</b> Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., Sugano S. Nat. Genet. 36:40-45(2004) [PubMed: 14702039] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA]. <u>Tissue:</u> Placenta.
[5]	NHLBI resequencing and genotyping service (RS&G) Submitted (DEC-2007) to the EMBL/GenBank/DDBJ databases <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA].
[6]	Mural R.J., Istrail S., Sutton G.G., Florea L., Halpern A.L., Mobarry C.M., Lippert R., Walenz B., Shatkay H., Dew I., Miller J.R., Flanigan M.J., Edwards N.J., Bolanos R., Fasulo D., Halldorsson B.V., Hannenhalli S., Turner R., Venter J.C. Submitted (SEP-2005) to the EMBL/GenBank/DDBJ databases <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
[7]	<b>"The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC)."</b> The MGC Project Team Genome Res. 14:2121-2127(2004) [PubMed: 15489334] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA]. <u>Tissue:</u> Brain.
[8]	<b>"Molecular cloning and amino acid sequence of human enkephalinase (neutral endopeptidase)."</b> Malfroy B., Kuang W.-J., Seeburg P.H., Mason A.J., Schofield P.R. FEBS Lett. 229:206-210(1988) [PubMed: 3162217] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] OF 3-750. <u>Tissue:</u> Placenta.
[9]	<b>"Endopeptidase-24.11 in human plasma degrades atrial natriuretic factor (ANF) to ANF (99-105/106-126)."</b> Yandle T.G., Brennan S.O., Espiner E.A., Nicholls M.G., Richards A.M. Peptides 10:891-894(1989) [PubMed: 2531377] [Abstract] <u>Cited for:</u> FUNCTION IN THE DEGRADATION OF ANF.
[10]	<b>"Asp650 is crucial for catalytic activity of neutral endopeptidase 24-11."</b> Le Moual H., Dion N., Roques B.P., Crine P., Boileau G. Eur. J. Biochem. 221:475-480(1994) [PubMed: 8168535] [Abstract] <u>Cited for:</u> ACTIVE SITE ASP-651.
[11]	<b>"Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry."</b>


	Zhang H., Li X.-J., Martin D.B., Aebersold R. Nat. Biotechnol. 21:660-666(2003) [PubMed: 12754519] [Abstract] <u>Cited for:</u> GLYCOSYLATION AT ASN-145 AND ASN-285.
[12]	<b>"Evaluation of angiotensin-converting enzyme (ACE), its homologue ACE2 and neprilysin in angiotensin peptide metabolism."</b> Rice G.I., Thomas D.A., Grant P.J., Turner A.J., Hooper N.M. Biochem. J. 383:45-51(2004) [PubMed: 15283675] [Abstract] <u>Cited for:</u> FUNCTION IN ANGIOTENSIN PEPTIDE METABOLISM, BIOPHYSICOCHEMICAL PROPERTIES.
[13]	<b>"Human opiorphin, a natural antinociceptive modulator of opioid-dependent pathways."</b> Wisner A., Dufour E., Messaoudi M., Nejd A., Marcel A., Ungeheuer M.-N., Rougeot C. Proc. Natl. Acad. Sci. U.S.A. 103:17979-17984(2006) [PubMed: 17101991] [Abstract] <u>Cited for:</u> INHIBITION BY OPIORPHIN.
[14]	<b>"Neprilysin is identical to skin fibroblast elastase: its role in skin aging and UV responses."</b> Morisaki N., Moriwaki S., Sugiyama-Nakagiri Y., Haketa K., Takema Y., Imokawa G. J. Biol. Chem. 285:39819-39827(2010) [PubMed: 20876573] [Abstract] <u>Cited for:</u> IDENTIFICATION AS SKIN FIBROBLAST ELASTASE, FUNCTION.
[15]	<b>"Structure of human neutral endopeptidase (Neprilysin) complexed with phosphoramidon."</b> Oefner C., D'Arcy A., Hennig M., Winkler F.K., Dale G.E. J. Mol. Biol. 296:341-349(2000) [PubMed: 10669592] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS); GLYCOSYLATION AT ASN-145; ASN-325 AND ASN-628.
[16]	<b>"Structural analysis of neprilysin with various specific and potent inhibitors."</b> Oefner C., Roques B.P., Fournie-Zaluski M.-C., Dale G.E. Acta Crystallogr. D 60:392-396(2004) [PubMed: 14747736] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS) OF 55-750 IN COMPLEXES WITH ZINC IONS AND SYNTHETIC INHIBITORS, DISULFIDE BONDS, COFACTOR, GLYCOSYLATION AT ASN-145; ASN-325 AND ASN-628.
[17]	<b>"Structural studies of a bifunctional inhibitor of neprilysin and DPP-IV."</b> Oefner C., Pierau S., Schulz H., Dale G.E. Acta Crystallogr. D 63:975-981(2007) [PubMed: 17704566] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS) OF 55-750 IN COMPLEX WITH ZINC IONS AND THE SYNTHETIC INHIBITOR MCB3937, DISULFIDE BONDS, COFACTOR, GLYCOSYLATION AT ASN-145; ASN-325 AND ASN-628.
+	Additional computationally mapped references.

## Web resources

Atlas of Genetics and Cytogenetics in Oncology and Haematology

## Cross-references

### Sequence databases

<ul style="list-style-type: none"> <li>⊗ EMBL</li> <li>⊗ GenBank</li> <li>⊗ DDBJ</li> </ul>	Y00811 mRNA. Translation: CAA68752.1. J03779 mRNA. Translation: AAA51915.1. M26628  M26627 Genomic DNA. Translation: AAA52294.1. AK291761 mRNA. Translation: BAF84450.1. EU326307 Genomic DNA. Translation: ACA05913.1.
---	--

	CH471052 Genomic DNA. Translation: EAW78754.1. CH471052 Genomic DNA. Translation: EAW78755.1. CH471052 Genomic DNA. Translation: EAW78756.1. CH471052 Genomic DNA. Translation: EAW78757.1. CH471052 Genomic DNA. Translation: EAW78758.1. BC101632 mRNA. Translation: AAI01633.1. BC101658 mRNA. Translation: AAI01659.1. X07166 mRNA. Translation: CAA30157.1. Different initiation.																																																						
IPI	IPI00247063.																																																						
PIR	HYHUN. A41387.																																																						
RefSeq	NP_000893.2. NM_000902.3. NP_009218.2. NM_007287.2. NP_009219.2. NM_007288.2. NP_009220.2. NM_007289.2.																																																						
UniGene	Hs.307734.																																																						
<b>3D structure databases</b>																																																							
<input checked="" type="radio"/> PDBe <input checked="" type="radio"/> RCSB PDB <input checked="" type="radio"/> PDBj	<table><tr><th>Entry</th><th>Method</th><th>Resolution (Å)</th><th>Chain</th><th>Positions</th><th>PDBsum</th></tr><tr><td>1DL9</td><td>model</td><td>-</td><td>A</td><td>508-750</td><td>[»]</td></tr><tr><td>1DMT</td><td>X-ray</td><td>2.10</td><td>A</td><td>55-750</td><td>[»]</td></tr><tr><td>1QVD</td><td>model</td><td>-</td><td>A</td><td>55-750</td><td>[»]</td></tr><tr><td>1R1H</td><td>X-ray</td><td>1.95</td><td>A</td><td>55-749</td><td>[»]</td></tr><tr><td>1R1I</td><td>X-ray</td><td>2.60</td><td>A</td><td>55-749</td><td>[»]</td></tr><tr><td>1R1J</td><td>X-ray</td><td>2.35</td><td>A</td><td>55-749</td><td>[»]</td></tr><tr><td>1Y8J</td><td>X-ray</td><td>2.25</td><td>A</td><td>55-749</td><td>[»]</td></tr><tr><td>2QPJ</td><td>X-ray</td><td>2.05</td><td>A</td><td>55-750</td><td>[»]</td></tr></table>	Entry	Method	Resolution (Å)	Chain	Positions	PDBsum	1DL9	model	-	A	508-750	[»]	1DMT	X-ray	2.10	A	55-750	[»]	1QVD	model	-	A	55-750	[»]	1R1H	X-ray	1.95	A	55-749	[»]	1R1I	X-ray	2.60	A	55-749	[»]	1R1J	X-ray	2.35	A	55-749	[»]	1Y8J	X-ray	2.25	A	55-749	[»]	2QPJ	X-ray	2.05	A	55-750	[»]
Entry	Method	Resolution (Å)	Chain	Positions	PDBsum																																																		
1DL9	model	-	A	508-750	[»]																																																		
1DMT	X-ray	2.10	A	55-750	[»]																																																		
1QVD	model	-	A	55-750	[»]																																																		
1R1H	X-ray	1.95	A	55-749	[»]																																																		
1R1I	X-ray	2.60	A	55-749	[»]																																																		
1R1J	X-ray	2.35	A	55-749	[»]																																																		
1Y8J	X-ray	2.25	A	55-749	[»]																																																		
2QPJ	X-ray	2.05	A	55-750	[»]																																																		
ProteinModelPortal	P08473.																																																						
SMR	P08473. Positions 55-750.																																																						
ModBase	Search...																																																						
<b>Protein-protein interaction databases</b>																																																							
IntAct	P08473. 74 interactions.																																																						
STRING	P08473.																																																						
<b>Protein family/group databases</b>																																																							
MEROPS	M13.001.																																																						
<b>PTM databases</b>																																																							
PhosphoSite	P08473.																																																						
<b>Proteomic databases</b>																																																							
PeptideAtlas	P08473.																																																						
PRIDE	P08473.																																																						
<b>Genome annotation databases</b>																																																							
Ensembl	ENST00000360490; ENSP00000353679; ENSG00000196549. ENST00000404362; ENSP00000384558; ENSG00000196549. ENST00000460393; ENSP00000418525; ENSG00000196549. ENST00000462745; ENSP00000419653; ENSG00000196549.																																																						

	ENST00000492661; ENSP00000420389; ENSG00000196549. ENST00000493237; ENSP00000417079; ENSG00000196549.
GeneID	4311.
KEGG	hsa:4311.
UCSC	uc003fab.1. human.
<b>Organism-specific databases</b>	
CTD	4311.
GeneCards	GC03P152191.
H-InvDB	HIX0003792.
HGNC	HGNC:7154. MME.
HPA	CAB000013.
MIM	120520. gene.
neXtProt	NX_P08473.
PharmGKB	PA30864.
GenAtlas	Search...
<b>Phylogenomic databases</b>	
eggNOG	prNOG07698.
GeneTree	ENSGT00550000074200.
HOGENOM	HBG701001.
HOVERGEN	HBG005554.
InParanoid	P08473.
OMA	TYRPEYA.
OrthoDB	EOG4XWFXB.
PhylomeDB	P08473.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.24.11. 247.
<b>Gene expression databases</b>	
ArrayExpress	P08473.
Bgee	P08473.
CleanEx	HS_MME.
Genevestigator	P08473.
GermOnline	ENSG00000196549. Homo sapiens.
<b>Family and domain databases</b>	
InterPro	IPR000718. Peptidase_M13. IPR018497. Peptidase_M13_C. IPR008753. Peptidase_M13_N. [Graphical view]
PANTHER	PTHR11733. Peptidase_M13. 1 hit.



Pfam	PF01431. Peptidase_M13. 1 hit. PF05649. Peptidase_M13_N. 1 hit. [Graphical view]
PRINTS	PR00786. NEPRILYSIN.
PROSITE	PS00142. ZINC_PROTEASE. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
DrugBank	DB00616. Candoxatril.
NextBio	16959.
SOURCE	Search...

**Entry information**

Entry name	NEP_HUMAN
Accession	Primary (citable) accession number: <b>P08473</b> Secondary accession number(s): A8K6U6, D3DNJ9, Q3MIX4
Entry history	Integrated into August 1, 1988 UniProtKB/Swiss-Prot: Last sequence update: January 23, 2007 Last modified: April 5, 2011 This is version 146 of the entry and version 2 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Disclaimer</b>	Any medical or genetic information present in this entry is provided for research, educational and informational purposes only. It is not in any way intended to be used as a substitute for professional medical advice, diagnosis, treatment or care.

**Relevant documents**

Human cell differentiation molecules  
CD nomenclature of surface proteins of human leucocytes and list of entries

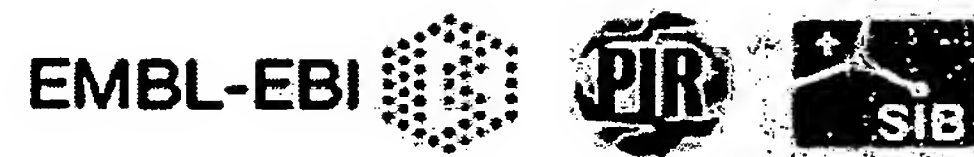
Human chromosome 3  
Human chromosome 3: entries, gene names and cross-references to MIM

MIM cross-references  
Online Mendelian Inheritance in Man (MIM) cross-references in UniProtKB/Swiss-Prot

PDB cross-references  
Index of Protein Data Bank (PDB) cross-references

Peptidase families  
Classification of peptidase families and list of entries

SIMILARITY comments  
Index of protein domains and families



## Protein

Translations of Life

Display Settings: GenPept

## common acute lymphoblastic leukemia antigen precursor [Homo sapiens]

GenBank: AAA51915.1

[FASTA](#) [Graphics](#)Go to:

LOCUS AAA51915 750 aa linear PRI 31-OCT-1994  
DEFINITION common acute lymphoblastic leukemia antigen precursor [Homo sapiens].  
ACCESSION AAA51915  
VERSION AAA51915.1 GI:179834  
DBSOURCE locus HUMCALLA accession [J03779.1](#)  
KEYWORDS .  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 750)  
AUTHORS Shipp,M.A., Richardson,N.E., Sayre,P.H., Brown,N.R., Masteller,E.L., Clayton,L.K., Ritz,J. and Reinherz,E.L.  
TITLE Molecular cloning of the common acute lymphoblastic leukemia antigen (CALLA) identifies a type II integral membrane protein  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 85 (13), 4819-4823 (1988)  
PUBMED [2968607](#)  
COMMENT Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..750  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="3q21-q27"  
Protein 1..750  
/name="common acute lymphoblastic leukemia antigen precursor"  
mat\_peptide 2..750  
/product="common acute lymphoblastic leukemia antigen"  
Region 72..750  
/region\_name="PepO"  
/note="Predicted metalloendopeptidase [Posttranslational modification, protein turnover, chaperones]; COG3590"  
/db\_xref="CDD:33390"  
Region 79..748  
/region\_name="M13"  
/note="Peptidase family M13 includes neprilysin, endothelin-converting enzyme I; cd08662"  
/db\_xref="CDD:189000"  
Site order(543..544,581,584..585,588,647,690..691,712,718)  
/site\_type="active"  
/db\_xref="CDD:189000"  
Site order(584,588,647)  
/site\_type="other"  
/note="Zn binding site"  
/db\_xref="CDD:189000"  
CDS 1..750  
/gene="MME"  
/coded\_by="J03779.1:12..2264"  
/db\_xref="GDB:G00-120-190"

ORIGIN  
1 mgksesqmdi tdintpkpkk kqrwtrleis lsvlvlllti iavrmialya tyddgickss  
61 dciksaarli qnmdattepc rdffkyacgg wlkrnvipet ssrygnfdil rdelevlkd  
121 vlqepktedi vavqkakaly rscinesaid srggepllk lpdigywpva tenweqkyga  
181 swtaekaiag lnskygkklv inlfvgtddk nsvnhvihid qprlgpsrd yyectgiyke  
241 actayvdfmi svarlirgee rlpidenqla lemknvmele keianatakp edrndpmilly  
301 nkmlraqiqn nlsleingkp fswlnftnei mstvnisitn eedvvvyape yltklkpilt  
361 kysardlqnl mswrfimdlv sslsrtykes rnafrkalyg ttsetatwrr canyvngnme  
421 navgrlyvea afageskhvv edliaqirev fiqtlldltw mdaetkkrae ekalaikeri  
481 gypddivsnd nklnneylel nykedeyfen iignlkfsqs kqlkkldrekv dkdewisgaa  
541 vvnafyssgr nqivfpagil qppffsaqgs nslnyggigm vigheithgf ddngrnfnkd  
601 gdlvdwtqq sasnfkeqsq cmvyqygnfs wdlaggqhl n gintlgenia dngglgqayr  
661 ayqnyikkng eekllpgldl nhkqlfflnf aqvwcgtyrp eyavnsiktd vhsppgnfrii  
721 gtlqnsaefts eafhcrknsy mnpekkcrvw

//

## Protein

Translations of Life

Display Settings: GenPept

## unnamed protein product [Homo sapiens]

GenBank: CAA68752.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS CAA68752 750 aa linear PRI 12-SEP-1993  
DEFINITION unnamed protein product [Homo sapiens].  
ACCESSION CAA68752  
VERSION CAA68752.1 GI:29626  
DBSOURCE embl accession Y00811.1  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Letarte,M., Vera,S., Tran,R., Addis,J.B., Onizuka,R.J.,  
Quackenbush,E.J., Jongeneel,C.V. and McInnes,R.R.  
TITLE Common acute lymphocytic leukemia antigen is identical to neutral  
endopeptidase  
JOURNAL J. Exp. Med. 168 (4), 1247-1253 (1988)  
PUBMED 2971756  
REFERENCE 2 (residues 1 to 750)  
AUTHORS Jongeneel,C.V.  
TITLE Direct Submission  
JOURNAL Submitted (11-AUG-1988) Jongeneel C.V., Ludwig Institute for Cancer  
Research, Lausanne Branch, Chemin des Boveresses, CH-1066  
Epalinges, Switzerland  
COMMENT Data kindly reviewed (31-AUG-1988) by Jongeneel C.V.  
FEATURES  
Location/Qualifiers  
source 1..750  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="45"  
/cell\_type="cortex"  
/tissue\_type="kidney"  
/clone\_lib="lambda gt11"  
Protein 1..750  
/name="unnamed protein product"  
Region 72..750  
/region\_name="PepO"  
/note="Predicted metalloendopeptidase [Posttranslational  
modification, protein turnover, chaperones]; COG3590"  
/db\_xref="CDD:33390"  
Region 79..748  
/region\_name="M13"  
/note="Peptidase family M13 includes neprilysin,  
endothelin-converting enzyme I; cd08662"  
/db\_xref="CDD:189000"  
Site order(543..544,581,584..585,588,647,690..691,712,718)  
/site\_type="active"  
/db\_xref="CDD:189000"  
Site order(584,588,647)  
/site\_type="other"  
/note="Zn binding site"  
/db\_xref="CDD:189000"  
CDS 1..750  
/coded\_by="Y00811.1:136..2388"  
/note="CALLA protein (AA 1 - 750)"  
/db\_xref="GDB:120190"  
/db\_xref="GOA:P08473"  
/db\_xref="HGNC:7154"  
/db\_xref="InterPro:IPR000718"  
/db\_xref="InterPro:IPR008753"  
/db\_xref="InterPro:IPR018497"  
/db\_xref="PDB:1DL9"  
/db\_xref="PDB:1DMT"  
/db\_xref="PDB:1QVD"  
/db\_xref="PDB:1R1H"  
/db\_xref="PDB:1R1I"  
/db\_xref="PDB:1R1J"  
/db\_xref="PDB:1Y8J"  
/db\_xref="PDB:2QPJ"  
/db\_xref="UniProtKB/Swiss-Prot:P08473"  
ORIGIN  
1 mgksesqmdi tdintpkpkk kqrwtpleis lsvlvlllti iavtmialya tyddgickss  
61 dciksaarli qnmdattepc tdfkkyacgg wlkrnvipet ssrygnfdil rdelevvlkd  
121 vlqepktedi vavqkakaly rscinesaid srggepllk1 lpdiygwpva tenweqkyga  
181 swtaekaiag lnskygkklv inlfvgtddk nsvnhvihid qprlglsard yyectgiyke  
241 actayvdfmi svarlirqee rlpidenqla lemknvmele keianatakp edrndpmlly  
301 nkmtlaqign nlsleingkp fswlnftnei mstvnisitn eedvvvyape yltklkpilt  
361 kysardlqnl mswrfimdlv sslsrtypes rnafrkalyg ttsetatwrr canyvngnme  
421 navgrlyvea afageskhvv edliaqirev fiqtlldltw mdaetkkrae ekalaikeri  
481 gypddivsnd nklnneylel nykedeyfen iiqnlfksqs kqlklrekv dkdewisgaa  
541 vvnafysgr nqivfpagil qppffsaqgs nslnyggigm vigheithgf ddngrnfnkd  
601 gdlvdwtqg sasnfkeqsg cmvygygnfs wdlaggqhlh gintlgenia dngglgqayr  
661 aygnyikkng eekllpgldl nhkqlfflnf agvwcgtyrp eyavnsiktd vhspgnfrii  
721 gtlqnsaefs eafhcrknsy mnpekkcrvw

//



## Protein

Translations of Life

Display Settings: GenPept

## enkephalinase [Homo sapiens]

GenBank: AAA52294.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAA52294 750 aa linear PRI 07-NOV-1994  
DEFINITION enkephalinase [Homo sapiens].  
ACCESSION AAA52294  
VERSION AAA52294.1 GI:179860  
DBSOURCE locus HUMCALLA03 accession M26607.1  
locus HUMCALLA04 accession M26608.1  
locus HUMCALLA05 accession M26609.1  
locus HUMCALLA06 accession M26610.1  
locus HUMCALLA07 accession M26611.1  
locus HUMCALLA08 accession M26612.1  
locus HUMCALLA09 accession M26613.1  
locus HUMCALLA10 accession M26614.1  
locus HUMCALLA11 accession M26615.1  
locus HUMCALLA12 accession M26616.1  
locus HUMCALLA13 accession M26617.1  
locus HUMCALLA14 accession M26618.1  
locus HUMCALLA15 accession M26619.1  
locus HUMCALLA16 accession M26620.1  
locus HUMCALLA17 accession M26621.1  
locus HUMCALLA18 accession M26622.1  
locus HUMCALLA19 accession M26623.1  
locus HUMCALLA20 accession M26624.1  
locus HUMCALLA21 accession M26625.1  
locus HUMCALLA22 accession M26626.1  
locus HUMCALLA23 accession M26627.1  
locus HUMCALLA24 accession M26628.1

KEYWORDS .  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 750)  
AUTHORS D'Adamio, L., Shipp, M. A., Masteller, E. L. and Reinherz, E. L.  
TITLE Organization of the gene encoding common acute lymphoblastic  
leukemia antigen (neutral endopeptidase 24.11): multiple minixons  
and separate 5' untranslated regions  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86 (18), 7103-7107 (1989)  
PUBMED 2528730  
COMMENT Method: conceptual translation.  
FEATURES  
source Location/Qualifiers  
1..750  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="3q21-q27"  
/cell\_line="NALM 6"  
Protein 1..750  
/product="enkephalinase"  
/EC\_number="3.4.24.11"  
Region 72..750  
/region\_name="PepO"  
/note="Predicted metalloendopeptidase [Posttranslational  
modification, protein turnover, chaperones]; COG3590"  
/db\_xref="CDD:33390"  
Region 79..748  
/region\_name="M13"  
/note="Peptidase family M13 includes neprilysin,  
endothelin-converting enzyme I; cd08662"  
/db\_xref="CDD:189000"  
Site order(543..544,581,584..585,588,647,690..691,712,718)  
/site\_type="active"  
/db\_xref="CDD:189000"  
Site order(584,588,647)  
/site\_type="other"  
/note="Zn binding site"  
/db\_xref="CDD:189000"  
CDS 1..750  
/gene="MME"  
/coded\_by="join(M26607.1:18..177,M26608.1:8..43,  
M26609.1:8..169,M26610.1:8..88,M26611.1:8..103,  
M26612.1:8..126,M26613.1:8..73,M26614.1:8..142,  
M26615.1:8..109,M26616.1:8..144,M26617.1:8..101,  
M26618.1:8..136,M26619.1:8..106,M26620.1:8..88,  
M26621.1:8..111,M26622.1:8..66,M26623.1:8..127,  
M26624.1:8..141,M26625.1:8..73,M26626.1:8..103,  
M26627.1:8..84,M26628.1:8..107)"  
/note="common acute lymphoblastic antigen"  
/db\_xref="GDB:G00-120-190"

ORIGIN  
1 mgksesqmdi tdintpkpkk kqrwtpleis lsvlvlllti iavtmialya tyddgickss  
61 dciksaarli qnmdattepc tdfkyacgg wlkrnvipet ssrygnfdil rdelevvlkd  
121 vlqepktedi vavqkakaly rscinesaid srggepllk1 lpdiygwpva tenweqkyga  
181 swtaekaiag lnskyykkvl inlfvgtddk nsnhvihid qprlgipsrd yyectgiyke  
241 actayvdfmi svarlirgee rlpidenqla lemnkvmele keianatakp edrndpmlly  
301 nkmtlaqiqn nlsleingkp fswlnftnei mstvnisitn eedvvvyape yltklkpilt  
361 kysardlqnl mswrfimdlv sslsrtykes rnafrkalyg ttsetatwrr canyvngnme  
421 navgrlyvea afageskhvv edliaqirev fiqtlldltw mdaetkkrae ekalaikeri  
481 gypddivsnd nklneyl1 nykedeyfen iignlkfsqs kqlklrekv dkdewisgaa

```

541 vvnafyssgr nqivfpagil qppffsaqqs nslnyggigm vigheithgf ddngrnfnkd
601 gdlvdwtqg sasnfeqsg cmvyqygns wdlaggqhl gintlgenia dngglgqayr
661 ayqnyikkng eekllpgldl nhkqlfflnf aqvwcgtyrp eyavnsiktd vhspgnfrii
721 gtlqnsaefs eafhcrknsy mpekkcrvw

```

//

## Protein

Translations of Life

Display Settings: GenPept

## unnamed protein product [Homo sapiens]

GenBank: CAA30157.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS CAA30157 743 aa linear PRI 12-SEP-1993  
DEFINITION unnamed protein product [Homo sapiens].  
ACCESSION CAA30157  
VERSION CAA30157.1 GI:34758  
DBSOURCE embl accession X07166.1  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 743)  
AUTHORS Malfroy,B., Kuang,W.J., Seeburg,P.H., Mason,A.J. and Schofield,P.R.  
TITLE Molecular cloning and amino acid sequence of human enkephalinase  
(neutral endopeptidase)  
JOURNAL FEBS Lett. 229 (1), 206-210 (1988)  
PUBMED 3162217  
FEATURES  
Location/Qualifiers  
source 1..743  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="lambda H7"  
/tissue\_type="placenta"  
/clone\_lib="lambda gt10"  
Protein 1..743  
/name="unnamed protein product"  
Region 65..743  
/region\_name="PepO"  
/note="Predicted metalloendopeptidase [Posttranslational  
modification, protein turnover, chaperones]; COG3590"  
/db\_xref="CDD:33390"  
Region 72..741  
/region\_name="M13"  
/note="Peptidase family M13 includes neprilysin,  
endothelin-converting enzyme I; cd08662"  
/db\_xref="CDD:189000"  
Site order(536..537,574,577..578,581,640,683..684,705,711)  
/site\_type="active"  
/db\_xref="CDD:189000"  
Site order(577,581,640)  
/site\_type="other"  
/note="Zn binding site"  
/db\_xref="CDD:189000"  
CDS 1..743  
/coded\_by="X07166.1:18..2249"  
/note="enkephalinase (AA 1-743) "  
/db\_xref="GDB:120190"  
/db\_xref="GOA:P08473"  
/db\_xref="HGNC:7154"  
/db\_xref="InterPro:IPR000718"  
/db\_xref="InterPro:IPR008753"  
/db\_xref="InterPro:IPR018497"  
/db\_xref="PDB:1DL9"  
/db\_xref="PDB:1DMT"  
/db\_xref="PDB:1QVD"  
/db\_xref="PDB:1R1H"  
/db\_xref="PDB:1R1I"  
/db\_xref="PDB:1R1J"  
/db\_xref="PDB:1Y8J"  
/db\_xref="PDB:2QPJ"  
/db\_xref="UniProtKB/Swiss-Prot:P08473"  
ORIGIN  
1 mditdintpk pkkkqrwtpl eislsvlvll ltiavtmia lyatyddgic kssdciksaa  
61 rliqnmhatt epctdffkya cggwlkrnvi petssrygnf dilrdelevv lkdvlpqekt  
121 edivavqkak alyrscines aidsrggepl lkllpdiygw pvaatenweqk ygaswtaeka  
181 iaqlnskygk kvlinlfvgt ddksnvnhvi hidqprlglp srddyectgi ykeactayvd  
241 fmsvarlir qeerpiden qlalemnkvm elekeianat akpedrndpm llynkmtlaq  
301 ignnfslein gkpfswnft neimstvnis itneedvvvy apeyltklkp iltkysardl  
361 qnlmswrfim dlvselsrty kesrnafrka lygttsetat wrrcanyvng nmenavgrly  
421 veaafagesk hvvedliaqi revfiqtldd ltwmdaetkk raekalaik erigypddiv  
481 sndnklney lellykedey feniignlkf sqskqlklr ekvdkdewis gaavvnafys  
541 sgrnqivfpa gilqppffsa qqsnslnygg igmvigheit hgfdngrnf nkdgdldvww  
601 tqqsasnfke qsqcmvyqyg nfwswlaggq hlngintlge niadngglgq arrayqnyik  
661 kngeekllpg ldlnhkqlff lnfaqvwcgt yrpeyavnsi ktdvhspgnf riigtlnsa  
721 efseafhcrk nsymnpekkc rvw  
//

**Q61391 (NEP\_MOUSE) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified March 8, 2011. Version 101.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Neprilysin</b> EC=3.4.24.11 <i>Alternative name(s):</i> Atriopeptidase Enkephalinase Neutral endopeptidase 24.11 Short name=NEP Short name=Neutral endopeptidase Skin fibroblast elastase Short name=SFE CD_antigen=CD10
Gene names	Name: <b>Mme</b>
Organism	<b>Mus musculus (Mouse)</b>
Taxonomic identifier	10090 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires › Rodentia › Sciurognathi › Muroidea › Muridae › Murinae › Mus › Mus

**Protein attributes**

Sequence length	750 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Thermolysin-like specificity, but is almost confined on acting on polypeptides of up to 30 amino acids. Biologically important in the destruction of opioid peptides such as Met- and Leu-enkephalins by cleavage of a Gly-Phe bond. Able to cleave angiotensin-1, angiotensin-2 and angiotensin 1-9 <a href="#">By similarity</a> . Involved in the degradation of atrial natriuretic factor (ANF). Displays UV-inducible elastase activity toward skin preelastic and elastic fibers. <a href="#">Ref.6</a>
Catalytic activity	Preferential cleavage of polypeptides between hydrophobic residues, particularly with Phe or Tyr at P1'.
Cofactor	Binds 1 zinc ion per subunit.
Subcellular location	Cell membrane; Single-pass type II membrane protein.
Sequence similarities	Belongs to the peptidase M13 family.



Ontologies

Keywords

Cellular component	Cell membrane Membrane
Domain	Signal-anchor Transmembrane Transmembrane helix
Ligand	Metal-binding Zinc
Molecular function	Hydrolase Metalloprotease Protease
PTM	Disulfide bond Glycoprotein
Technical term	3D-structure Direct protein sequencing

Gene Ontology (GO)

Cellular component	integral to membrane Inferred from electronic annotation. Source: UniProtKB-KW plasma membrane Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	metal ion binding Inferred from electronic annotation. Source: UniProtKB-KW metalloendopeptidase activity Inferred from electronic annotation. Source: InterPro

Complete GO annotation...

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
Molecule processing					
<input type="checkbox"/> Initiator methionine	1	1	Removed By similarity	I	
<input checked="" type="checkbox"/> Chain	2 – 750	749	Neprilysin		PRO_0000078
Regions					
<input checked="" type="checkbox"/> Topological domain	2 – 28	27	Cytoplasmic Potential	I	
<input checked="" type="checkbox"/> Transmembrane	29 – 51	23	Helical; Signal-anchor for type II	II	

<input checked="" type="checkbox"/>	Topological domain	52 – 750	699	membrane protein; Potential Extracellular Potential		
<input type="checkbox"/>	Motif	16 – 23	8	Stop-transfer sequence Potential		
Sites						
<input type="checkbox"/>	Active site	585	1	By similarity		
<input type="checkbox"/>	Active site	651	1	Proton donor By similarity		
<input type="checkbox"/>	Metal binding	584	1	Zinc; catalytic By similarity		
<input type="checkbox"/>	Metal binding	588	1	Zinc; catalytic By similarity		
<input type="checkbox"/>	Metal binding	647	1	Zinc; catalytic By similarity		
<input type="checkbox"/>	Binding site	103	1	Substrate carboxyl By similarity		
Amino acid modifications						
<input type="checkbox"/>	Glycosylation	145	1	N-linked (GlcNAc...) Ref.5		
<input type="checkbox"/>	Glycosylation	211	1	N-linked (GlcNAc...) Potential		
<input type="checkbox"/>	Glycosylation	285	1	N-linked (GlcNAc...) Ref.5		
<input type="checkbox"/>	Glycosylation	311	1	N-linked (GlcNAc...) Ref.5		
<input type="checkbox"/>	Glycosylation	317	1	N-linked (GlcNAc...) Ref.5		
<input type="checkbox"/>	Glycosylation	325	1	N-linked (GlcNAc...) By similarity		

<input type="checkbox"/>	Glycosylation	628	1	N-linked (GlcNAc...) By similarity		
<input type="checkbox"/>	Disulfide bond	57 ↔ 62		By similarity		
<input type="checkbox"/>	Disulfide bond	80 ↔ 735		By similarity		
<input type="checkbox"/>	Disulfide bond	88 ↔ 695		By similarity		
<input type="checkbox"/>	Disulfide bond	143 ↔ 411		By similarity		
<input type="checkbox"/>	Disulfide bond	234 ↔ 242		By similarity		
<input type="checkbox"/>	Disulfide bond	621 ↔ 747		By similarity		

Experimental info

<input type="checkbox"/>	Sequence conflict	230	1	D → G in AAA37386. Ref.1		
--------------------------	-------------------	-----	---	--------------------------------	--	--

Secondary structure

1 ...

■ Helix ■ Strand ■ Turn

Details...

Sequences

Sequence	Length	Mass (Da)
<div><input checked="" type="checkbox"/> Q61391 [UniParc]. Last modified January 23, 2007. Version 3. Checksum: 1FC39A971D98F6FE</div>	FASTA 750	85,702
<div><div><div><div>102030405060</div><div>MGRSESQMDITDINAPKPKKQQRWTPLEISLSVLVLLLTIIAVTMIALYA TYDDGICKSS</div></div><div><div>708090100110120</div><div>DCIKSAARLI QNMDASVEPC TDFFKYACGG WLKRNVIPET SSRYSNFDIL RDELEVILKD</div></div><div><div>130140150160170180</div><div>VLQEPKTEDI VAVQKAKTLY RSCINESAID SRGGQPLLKL LPDIYGWPVA SDNWDQTYGT</div></div><div><div>190200210220230240</div><div>SWTAEKSIAQ LNSKYGKKVL INFFVGTDDK NSTQHIHFDPRLGLPSRD YYECTGIYKE</div></div><div><div>250260270280290300</div><div>ACTAYVDFMI SVALIRQEQ SLPIDENQLS LEMNKVMELE KEIANATTKP EDRNDPMLLY</div></div><div><div>310320330340350360</div><div>NKMTLAKLQN NFSLEVNGKS FSWSNFTNEI MSTVNINIQN EEEVVVYAPE YLTKLKPILT</div></div><div><div>370380390400410420</div><div>KYSPRDLQNL MSWRFIMDLV SLSRNYKES RNAFRKALYG TTSETATWRR CANYVNGNME</div></div><div><div>430440450460470480</div><div>NAVGRLYVEA AFAGESKHVV EDLIAQIREV FIQTLDDLTW MDAETKKKAE EKALAIKERI</div></div><div><div>490500510520530540</div><div>GYPDDIISNE NKLNNEYLEL NYREDEYFEN IIQNLKFSQS KQLKKLREKV DKDEWISGAA</div></div><div><div>550560570580590600</div><div>VVNAFYSSGR NQIVFPAGIL QPPFFSAQQS NSLNYGGIGM VIGHEITHGF DDNGRNFNKD</div></div><div><div>610620630640650660</div><div></div></div></div></div>		

```

GDLVDWWTQQ SANNFKDQSQ CMVYQYGNFS WDLAGGQHLN GINTLGENIA DNGGIGQAYR
      670      680      690      700      710      720
AYQNYVKKNG EEKLLPGLDL NHKQLFFLNF AQVWCGTYRP EYAVNSIKTD VHSPGNFRII
      730      740      750
GTLQNSAEFA DAFHCRKNSY MNPERKCRVW

```

« Hide

## References

« Hide 'large scale' references

- [1] **"Murine common acute lymphoblastic leukemia antigen (CD10 neutral endopeptidase 24.11). Molecular characterization, chromosomal localization, and modeling of the active site."**  
Chen C.Y., Salles G., Seldin M.F., Kister A.E., Reinher E.L., Shipp M.A.  
J. Immunol. 148:2817-2825(1992) [PubMed: 1374101] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA].  
Strain: BALB/c.
- [2] **"The transcriptional landscape of the mammalian genome."**  
Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N., Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K., Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M., Davis M.J., Hayashizaki Y.  
Science 309:1559-1563(2005) [PubMed: 16141072] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
Strain: C57BL/6J.  
Tissue: Epididymis and Testis.
- [3] **"The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC)."**  
The MGC Project Team  
Genome Res. 14:2121-2127(2004) [PubMed: 15489334] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
Strain: C57BL/6J and FVB/N.  
Tissue: Embryonic germ cell and Mammary tumor.
- [4] Lubec G., Sunyer B., Chen W.-Q.  
Submitted (JAN-2009) to UniProtKB  
Cited for: PROTEIN SEQUENCE OF 473-479, MASS SPECTROMETRY.  
Strain: OF1.  
Tissue: Hippocampus.
- [5] **"Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins."**  
Wollscheid B., Bausch-Fluck D., Henderson C., O'Brien R., Bibel M., Schiess R., Aebersold R., Watts J.D.  
Nat. Biotechnol. 27:378-386(2009) [PubMed: 19349973] [Abstract]  
Cited for: GLYCOSYLATION [LARGE SCALE ANALYSIS] AT ASN-145; ASN-285; ASN-311 AND ASN-317, MASS SPECTROMETRY.
- [6] **"Neprilysin is identical to skin fibroblast elastase: its role in skin aging and UV responses."**  
Morisaki N., Moriwaki S., Sugiyama-Nakagiri Y., Haketa K., Takema Y., Imokawa G.  
J. Biol. Chem. 285:39819-39827(2010) [PubMed: 20876573] [Abstract]  
Cited for: IDENTIFICATION AS SKIN FIBROBLAST ELASTASE, FUNCTION.
- + Additional computationally mapped references.





**Phylogenomic databases**

eggNOG	roNOG08854.
HOGENOM	HBG701001.
HOVERGEN	HBG005554.
InParanoid	Q61391.
OMA	TYRPEYA.
OrthoDB	EOG4XWFXB.
PhylomeDB	Q61391.

**Enzyme and pathway databases**

BRENDA	3.4.24.11. 244.
--------	-----------------

**Gene expression databases**

ArrayExpress	Q61391.
Bgee	Q61391.
CleanEx	MM_MME.
Genevestigator	Q61391.
GermOnline	ENSMUSG00000027820. Mus musculus.

**Family and domain databases**

InterPro	IPR000718. Peptidase_M13. IPR018497. Peptidase_M13_C. IPR008753. Peptidase_M13_N. [Graphical view]
PANTHER	PTHR11733. Peptidase_M13. 1 hit.
Pfam	PF01431. Peptidase_M13. 1 hit. PF05649. Peptidase_M13_N. 1 hit. [Graphical view]
PRINTS	PR00786. NEPRILYSIN.
PROSITE	PS00142. ZINC_PROTEASE. 1 hit. [Graphical view]
ProtoNet	Search...

**Other Resources**

NextBio	291980.
SOURCE	Search...

**Entry information**

Entry name	NEP_MOUSE
Accession	Primary (citable) accession number: <b>Q61391</b> Secondary accession number(s): Q6NXX5, Q8K251
Entry history	Integrated into November 1, 1997 UniProtKB/Swiss- Prot:

	Last sequence update: January 23, 2007 Last modified: March 8, 2011 This is version 101 of the entry and version 3 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Relevant documents</b>	
MGD cross-references Mouse Genome Database (MGD) cross-references in UniProtKB/Swiss-Prot	
PDB cross-references Index of Protein Data Bank (PDB) cross-references	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



**P08049 (NEP\_RABIT) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified March 8, 2011. Version 95.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Neprilysin</b> EC=3.4.24.11 <i>Alternative name(s):</i> Atriopeptidase Enkephalinase Neutral endopeptidase 24.11 Short name=NEP Short name=Neutral endopeptidase Skin fibroblast elastase Short name=SFE CD_antigen=CD10
Gene names	Name: <b>MME</b>
Organism	<b>Oryctolagus cuniculus (Rabbit)</b>
Taxonomic identifier	9986 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires › Lagomorpha › Leporidae › Oryctolagus

**Protein attributes**

Sequence length	750 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Thermolysin-like specificity, but is almost confined on acting on polypeptides of up to 30 amino acids. Biologically important in the destruction of opioid peptides such as Met- and Leu-enkephalins by cleavage of a Gly-Phe bond. Able to cleave angiotensin-1, angiotensin-2 and angiotensin 1-9 <span>By similarity</span> . Involved in the degradation of atrial natriuretic factor (ANF). Displays UV-inducible elastase activity toward skin preelastic and elastic fibers <span>By similarity</span> .
Catalytic activity	Preferential cleavage of polypeptides between hydrophobic residues, particularly with Phe or Tyr at P1'.
Cofactor	Binds 1 zinc ion per subunit.
Subcellular location	Cell membrane; Single-pass type II membrane protein.
Sequence similarities	Belongs to the peptidase M13 family.



## Ontologies

### Keywords

Cellular component	Cell membrane Membrane
Domain	Signal-anchor Transmembrane Transmembrane helix
Ligand	Metal-binding Zinc
Molecular function	Hydrolase Metalloprotease Protease
PTM	Disulfide bond Glycoprotein
Technical term	Direct protein sequencing

### Gene Ontology (GO)

Biological process	proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	integral to membrane Inferred from electronic annotation. Source: UniProtKB-KW plasma membrane Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	metal ion binding Inferred from electronic annotation. Source: UniProtKB-KW metalloendopeptidase activity Inferred from electronic annotation. Source: InterPro

Complete GO annotation...

## Sequence annotation (Features)

	Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
<b>Molecule processing</b>						
<input type="checkbox"/>	Initiator methionine	1	1	Removed	I	
<input checked="" type="checkbox"/>	Chain	2 – 750	749	Neprilysin		PRO_0000078
<b>Regions</b>						
<input checked="" type="checkbox"/>	Topological domain	2 – 28	27	Cytoplasmic Potential	I	
<input checked="" type="checkbox"/>	Transmembrane	29 – 51	23	Helical; Signal- anchor for	II	

<input type="checkbox"/>	Topological domain	52 – 750	699	type II membrane protein; Potential		
				Extracellular Potential		
<input type="checkbox"/>	Motif	16 – 23	8	Stop-transfer sequence Potential		
<b>Sites</b>						
<input type="checkbox"/>	Active site	585	1	By similarity		
<input type="checkbox"/>	Active site	651	1	Proton donor By similarity		
<input type="checkbox"/>	Metal binding	584	1	Zinc; catalytic		
<input type="checkbox"/>	Metal binding	588	1	Zinc; catalytic		
<input type="checkbox"/>	Metal binding	647	1	Zinc; catalytic By similarity		
<input type="checkbox"/>	Binding site	103	1	Substrate carboxyl By similarity		
<b>Amino acid modifications</b>						
<input type="checkbox"/>	Glycosylation	145	1	N-linked (GlcNAc...) By similarity		
<input type="checkbox"/>	Glycosylation	285	1	N-linked (GlcNAc...) Potential		
<input type="checkbox"/>	Glycosylation	311	1	N-linked (GlcNAc...) By similarity		
<input type="checkbox"/>	Glycosylation	325	1	N-linked (GlcNAc...) By similarity		
<input type="checkbox"/>	Glycosylation	628	1	N-linked (GlcNAc...)		
<input type="checkbox"/>	Disulfide bond	57 ↔ 62		By similarity		
<input type="checkbox"/>	Disulfide bond	80 ↔ 735		By similarity		
<input type="checkbox"/>	Disulfide bond	88 ↔ 695		By similarity		
<input type="checkbox"/>	Disulfide bond	143 ↔ 411		By similarity		
<input type="checkbox"/>	Disulfide bond	234 ↔ 242		By similarity		

☐ Disulfide bond

621 ↔ 747

☐ By similarity

☐

☐

Sequences

Sequence	Length	Mass (Da)
<input type="checkbox"/> P08049 [UniParc]. Last modified January 23, 2007. Version 2. Checksum: 0F26AF7316BB9D12	FASTA 750	85,582

10	20	30	40	50	60
MGRSESQMDI	TDINTPKPKK	KQRWTPLEIS	LSVLVLLLTV	IAVTMIALYA	TYDDGICKSS
70	80	90	100	110	120
DCIKSAARLI	QNMDATAEPC	TDFFKYACGG	WLKRNVIPET	SSRYSNFDIL	RDELEVILKD
130	140	150	160	170	180
VLQEPKTEDI	VAVQKAKTLY	RSCVNETAID	SRGGQPLLKL	LPDVYGWPVA	TQNWEQTYGT
190	200	210	220	230	240
SWSAEKSIAQ	LNSNYGKKVL	INFFVGTTDDK	NSMNHIIHID	QPRGLGPSRD	YYECTGIYKE
250	260	270	280	290	300
ACTAYVDFMI	AVAKLIRQEE	GLPIDENQIS	VEMNKVMELE	KEIANATTKS	EDRNDPMLLY
310	320	330	340	350	360
NKMTLAQIQN	NFSLEINGKP	FSWSNFTNEI	MSTVNINIPN	EEDVVVYAPF	YLIKLPILT
370	380	390	400	410	420
KYFPRDFQNL	FSWRFIMDLV	SSLSRTYKDS	RNAFRKALYG	TTSESATWRR	CANYVNGNME
430	440	450	460	470	480
NAVGRLYVEA	AFAGESKHVV	EDLIAQIREV	FIQTLDDLTV	MDAETKKKAE	EKALAIKERI
490	500	510	520	530	540
GYPDDIVSND	NKLNNEYLEL	NYKEDEYFEN	IIQNLKFSQS	KQLKKLREKV	DKDEWITGAA
550	560	570	580	590	600
IVNAFYSSGR	NQIVFPAGIL	QPPFFSAQQS	NSLNYGGIGM	VIGHEITHGF	DDNGRNFNKD
610	620	630	640	650	660
GDLVDWWTQQ	SANNFKEQSQ	CMVYQYGNFS	WDLAGGQHLN	GINTLGENIA	DNGGIGQAYR
670	680	690	700	710	720
AYQNYVKKNG	EEKLLPGIDL	NHKQLFFLNF	AQVWCGTYRP	EYAVNSIKTD	VHSPGNFRII
730	740	750			
GSLQNSVEFS	EAFCQPKNSY	MNPEKKCRVW			

« Hide

References

[1]

**"Amino acid sequence of rabbit kidney neutral endopeptidase 24.11 (enkephalinase) deduced from a complementary DNA."**  
Devault A., Lazure C., Nault C., le Moual H., Seidah N.G., Chretien M., Kahn P., Powell J., Mallet J., Beaumont A., Roques B.P., Crine P., Boileau G.  
EMBO J. 6:1317-1322(1987) [PubMed: 2440677] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA], PARTIAL PROTEIN SEQUENCE.  
Tissue: Kidney.

[2]

**Erratum**  
Devault A., Lazure C., Nault C., le Moual H., Seidah N.G., Chretien M., Kahn P., Powell J., Mallet J., Beaumont A., Roques B.P., Crine P., Boileau G.  
EMBO J. 6:2506-2506(1987)

- [3] **"An antibody purified with a lambda GT11 fusion protein precipitates enkephalinase activity."**  
Kahn P.H., Powell J.F., Beaumont A., Roques B.P., Mallet J.J.  
Biochem. Biophys. Res. Commun. 145:488-493(1987) [PubMed: 3297057] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA] OF 206-274.
- [4] **"Exploration of the catalytic site of endopeptidase 24.11 by site-directed mutagenesis. Histidine residues 583 and 587 are essential for catalysis."**  
Devault A., Sales V., Nault C., Beaumont A., Roques B., Crine P., Boileau G.  
FEBS Lett. 231:54-58(1988) [PubMed: 3162886] [Abstract]  
Cited for: MUTAGENESIS.

## Cross-references

### Sequence databases

<ul style="list-style-type: none"> <li>⊕ EMBL</li> <li>⊕ GenBank</li> <li>⊕ DDBJ</li> </ul>	X05338 mRNA. Translation: CAA28950.1. M16593 mRNA. Translation: AAA53694.1.
---	--

PIR	HYRBN. A29451.
-----	----------------

RefSeq	NP_001095155.1. NM_001101685.1.
--------	---------------------------------

UniGene	Ocu.2011.
---------	-----------

### 3D structure databases

ProteinModelPortal	P08049.
--------------------	---------

SMR	P08049. Positions 55-750.
-----	---------------------------

ModBase	Search...
---------	-----------

### Protein-protein interaction databases

STRING	P08049.
--------	---------

### Protein family/group databases

MEROPS	M13.001.
--------	----------

### Genome annotation databases

Ensembl	ENSOCUT00000017925; ENSOCUP00000015401; ENSOCUG00000017924.
---------	--

GeneID	100009251.
--------	------------

### Organism-specific databases

CTD	100009251.
-----	------------

### Phylogenomic databases

eggNOG	maNOG15085.
--------	-------------

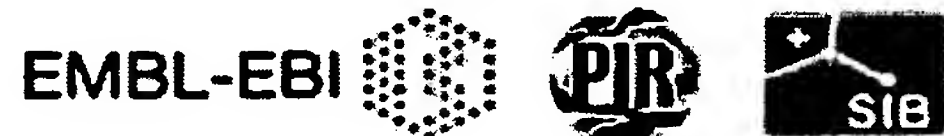
GeneTree	ENSGT00550000074200.
----------	----------------------

HOVERGEN	HBG005554.
----------	------------



OrthoDB	EOG4XWFXB.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.24.11. 255.
<b>Family and domain databases</b>	
InterPro	IPR000718. Peptidase_M13. IPR018497. Peptidase_M13_C. IPR008753. Peptidase_M13_N. [Graphical view]
PANTHER	PTHR11733. Peptidase_M13. 1 hit.
Pfam	PF01431. Peptidase_M13. 1 hit. PF05649. Peptidase_M13_N. 1 hit. [Graphical view]
PRINTS	PR00786. NEPRILYSIN.
PROSITE	PS00142. ZINC_PROTEASE. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Entry information</b>	
Entry name	NEP_RABIT
Accession	Primary (citable) accession number: <b>P08049</b>
Entry history	Integrated into August 1, 1988 UniProtKB/Swiss-Prot: Last sequence January 23, 2007 update: Last modified: March 8, 2011 This is version 95 of the entry and version 2 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Relevant documents</b>	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License &amp; Disclaimer | Contact



**P07861 (NEP\_RAT) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified March 8, 2011. Version 106.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Neprilysin</b> EC=3.4.24.11 <i>Alternative name(s):</i> Atriopeptidase Enkephalinase Neutral endopeptidase 24.11 Short name=NEP Short name=Neutral endopeptidase Skin fibroblast elastase Short name=SFE CD_antigen=CD10
Gene names	Name:Mme
Organism	Rattus norvegicus (Rat)
Taxonomic identifier	10116 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires › Rodentia › Sciurognathi › Muroidea › Muridae › Murinae › Rattus

**Protein attributes**

Sequence length	750 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Thermolysin-like specificity, but is almost confined on acting on polypeptides of up to 30 amino acids. Biologically important in the destruction of opioid peptides such as Met- and Leu-enkephalins by cleavage of a Gly-Phe bond. Able to cleave angiotensin-1, angiotensin-2 and angiotensin 1-9. Displays UV-inducible elastase activity toward skin preelastic and elastic fibers <span>By similarity</span> . Involved in the degradation of atrial natriuretic factor (ANF).
Catalytic activity	Preferential cleavage of polypeptides between hydrophobic residues, particularly with Phe or Tyr at P1'.
Cofactor	Binds 1 zinc ion per subunit.
Enzyme regulation	Inhibited in a dose dependent manner by sialorphin.
Subcellular location	Cell membrane; Single-pass type II membrane protein.

Sequence similarities		Belongs to the peptidase M13 family.																		
Ontologies																				
Keywords																				
Cellular component		Cell membrane Membrane																		
Domain		Signal-anchor Transmembrane Transmembrane helix																		
Ligand		Metal-binding Zinc																		
Molecular function		Hydrolase Metalloprotease Protease																		
PTM		Disulfide bond Glycoprotein																		
Technical term		Direct protein sequencing																		
Gene Ontology (GO)																				
Biological process		peptide metabolic process Inferred from mutant phenotype. Source: RGD  proteolysis Inferred from direct assay. Source: RGD																		
Cellular component		integral to membrane Inferred from electronic annotation. Source: UniProtKB-KW  membrane fraction Inferred from direct assay. Source: RGD  plasma membrane Inferred from mutant phenotype. Source: RGD																		
Molecular function		metal ion binding Inferred from electronic annotation. Source: UniProtKB-KW  metalloendopeptidase activity Inferred from electronic annotation. Source: InterPro  peptide binding Inferred from direct assay. Source: RGD																		
Complete GO annotation...																				
Sequence annotation (Features)																				
<table><tr><td>Feature key</td><td>Position(s)</td><td>Length</td><td>Description</td><td>Graphical view</td><td colspan="2">Feature identifier</td></tr><tr><td><input type="checkbox"/></td><td>Initiator methionine</td><td>1</td><td>1</td><td>Removed</td><td colspan="2"> </td></tr></table>							Feature key	Position(s)	Length	Description	Graphical view	Feature identifier		<input type="checkbox"/>	Initiator methionine	1	1	Removed		
Feature key	Position(s)	Length	Description	Graphical view	Feature identifier															
<input type="checkbox"/>	Initiator methionine	1	1	Removed																
Molecule processing																				
<table><tr><td><input type="checkbox"/></td><td>Initiator methionine</td><td>1</td><td>1</td><td>Removed</td><td colspan="2"> </td></tr></table>							<input type="checkbox"/>	Initiator methionine	1	1	Removed									
<input type="checkbox"/>	Initiator methionine	1	1	Removed																

<input checked="" type="checkbox"/>	Chain	2 – 750	749	Neprilysin		PRO_0000078
<b>Regions</b>						
<input checked="" type="checkbox"/>	Topological domain	2 – 28	27	Cytoplasmic <small>Potential</small>	I	
<input checked="" type="checkbox"/>	Transmembrane	29 – 51	23	Helical; Signal- anchor for type II membrane protein; <small>Potential</small>	II	
<input checked="" type="checkbox"/>	Topological domain	52 – 750	699	Extracellular <small>Potential</small>	I	
<input type="checkbox"/>	Motif	16 – 23	8	Stop- transfer sequence <small>Potential</small>	I	
<b>Sites</b>						
<input type="checkbox"/>	Active site	585	1	<small>By similarity</small>	I	
<input type="checkbox"/>	Active site	651	1	Proton donor <small>By similarity</small>	I	
<input type="checkbox"/>	Metal binding	584	1	Zinc; catalytic <small>By similarity</small>	I	
<input type="checkbox"/>	Metal binding	588	1	Zinc; catalytic <small>By similarity</small>	I	
<input type="checkbox"/>	Metal binding	647	1	Zinc; catalytic <small>By similarity</small>	I	
<input type="checkbox"/>	Binding site	103	1	Substrate carboxyl	I	
<b>Amino acid modifications</b>						
<input type="checkbox"/>	Glycosylation	145	1	N-linked (GlcNAc...) <small>By similarity</small>	I	
<input type="checkbox"/>	Glycosylation	211	1	N-linked (GlcNAc...) <small>Potential</small>	I	
<input type="checkbox"/>	Glycosylation	285	1	N-linked (GlcNAc...) <small>Potential</small>	I	
<input type="checkbox"/>	Glycosylation	311	1	N-linked (GlcNAc...) <small>Ref.1</small>	I	



<input type="checkbox"/>	Glycosylation	325	1	N-linked (GlcNAc...) By similarity		
<input type="checkbox"/>	Glycosylation	628	1	N-linked (GlcNAc...) By similarity		
<input type="checkbox"/>	Disulfide bond	57 ↔ 62		By similarity		
<input type="checkbox"/>	Disulfide bond	80 ↔ 735		By similarity		
<input type="checkbox"/>	Disulfide bond	88 ↔ 695		By similarity		
<input type="checkbox"/>	Disulfide bond	143 ↔ 411		By similarity		
<input type="checkbox"/>	Disulfide bond	234 ↔ 242		By similarity		
<input type="checkbox"/>	Disulfide bond	621 ↔ 747		By similarity		

Sequences

Sequence	Length	Mass (Da)
<div><input type="checkbox"/> P07861 [UniParc]. Last modified January 23, 2007. Version 2. Checksum: 609D4716C4B15CBC</div>	FASTA	75085,795
<div><div>102030405060</div><div>MGRSESQMDI TDINAPKPKK KQRWTPLEIS LSVLVLLLT IAVTMIALYA TYDDGICKSS</div><div>708090100110120</div><div>DCIKSAARLI QNMDASAEP TDFFKYACGG WLKRNVIPET SSRYSNFDIL RDELEVILKD</div><div>130140150160170180</div><div>VLQEPKTEDI VAVQKAKTLY RSCINESAID SRGGQPLLTL LPDIYGWPVA SQNWEQTYGT</div><div>190200210220230240</div><div>SWTAEKSIAQ LNSKYGKKVL INFFVGTDDK NSTQHIIHFD QPRLGLPSRD YYECTGIYKE</div><div>250260270280290300</div><div>ACTAYVDFMI SVARLIRQEQ RLPIDENQLS LEMNKVMELE KEIANATTKP EDRNDPMLLY</div><div>310320330340350360</div><div>NKMTLAKLQN NFSLEINGK FWSNFTNEI MSTVNINIQN EEEVVVYAPE YLTKLKPILT</div><div>370380390400410420</div><div>KYSPRDLQNL MSWRFIMDLV SSLSRNYKES RNAFRKALYG TTSETATWRR CANYVNGNME</div><div>430440450460470480</div><div>NAVGRLYVEA AFAGESKHVV EDLIAQIREV FIQTLDDL TW MDAETKKKAE EKALAIKERI</div><div>490500510520530540</div><div>GYPDDIISNE NKLNNEYLEL NYKEEEYFEN IIQNLKFSQS KQLKKLREKV DKDEWISGAA</div><div>550560570580590600</div><div>VVNAFYSSGR NQIVFPAGIL QPPFFSARQS NSLNYGGIGM VIGHEITHGF DDNGRNFNKD</div><div>610620630640650660</div><div>GDLVDWWTQQ SANNFKDQSQ CMVYQYGNFT WDLAGGQHLN GINTLGENIA DNGGIGQAYR</div><div>670680690700710720</div><div>AYQNYVKKNG EEKLLPGLDL NHKQLFFLNF AQVWCGTYRP EYAVNSIKTD VHSPGNFRII</div><div>730740750</div><div>GTLQNSAEFA DAFHCRKNSY MNPERKCRVW</div></div>		

« Hide

## References

	« Hide 'large scale' references
[1]	<p><b>"Molecular cloning and amino acid sequence of rat enkephalinase."</b>  Malfroy B., Schofield P.R., Kuang W.-J., Seeburg P.H., Mason A.J., Henzel W.J.  Biochem. Biophys. Res. Commun. 144:59-66(1987) [PubMed: 3555489] [Abstract]  Cited for: NUCLEOTIDE SEQUENCE [MRNA], PARTIAL PROTEIN SEQUENCE.</p>
[2]	<p><b>"The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC)."</b>  The MGC Project Team  Genome Res. 14:2121-2127(2004) [PubMed: 15489334] [Abstract]  Cited for: NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  Tissue: Kidney.</p>
[3]	<p><b>"Identification of the active-site arginine in rat neutral endopeptidase 24.11 (enkephalinase) as arginine 102 and analysis of a glutamine 102 mutant."</b>  Bateman R.C. Jr., Jackson D., Slaughter C.A., Unnithan S., Chai Y.G., Moomaw C., Hersh L.B.  J. Biol. Chem. 264:6151-6157(1989) [PubMed: 2703483] [Abstract]  Cited for: ACTIVE SITE ARG-103.</p>
[4]	<p><b>"Sialorphin, a natural inhibitor of rat membrane-bound neutral endopeptidase that displays analgesic activity."</b>  Rougeot C., Messaoudi M., Hermitte V., Rigault A.G., Blisnick T., Dugave C., Desor D., Rougeon F.  Proc. Natl. Acad. Sci. U.S.A. 100:8549-8554(2003) [PubMed: 12835417] [Abstract]  Cited for: INHIBITION BY SIALORPHIN.</p>
+	Additional computationally mapped references.

## Cross-references

### Sequence databases

<input checked="" type="radio"/> EMBL <input checked="" type="radio"/> GenBank <input checked="" type="radio"/> DDBJ	M15944 mRNA. Translation: AAA41116.1. BC085753 mRNA. Translation: AAH85753.1.
IPI	IPI00231789.
PIR	HYRTN. A29295.
RefSeq	NP_036740.1. NM_012608.2.
UniGene	Rn.33598.

### 3D structure databases

ProteinModelPortal	P07861.
SMR	P07861. Positions 55-750.
ModBase	Search...

### Protein-protein interaction databases

STRING	P07861.
--------	---------

**Protein family/group databases**

MEROPS	M13.001.
--------	----------

**Proteomic databases**

PRIDE	P07861.
-------	---------

**Genome annotation databases**

Ensembl	ENSRNOT00000042576; ENSRNOP00000044578; ENSRNOG00000009514.
---------	---

GeneID	24590.
--------	--------

KEGG	rno:24590.
------	------------

UCSC	NM_012608. rat.
------	-----------------

**Organism-specific databases**

CTD	24590.
-----	--------

RGD	3098. Mme.
-----	------------

**Phylogenomic databases**

eggNOG	roNOG08854.
--------	-------------

GeneTree	ENSGT00550000074200.
----------	----------------------

HOVERGEN	HBG005554.
----------	------------

InParanoid	P07861.
------------	---------

OMA	TYRPEYA.
-----	----------

OrthoDB	EOG4XWFXB.
---------	------------

PhylomeDB	P07861.
-----------	---------

**Enzyme and pathway databases**

BRENDA	3.4.24.11. 248.
--------	-----------------

**Gene expression databases**

ArrayExpress	P07861.
--------------	---------

Genevestigator	P07861.
----------------	---------

GermOnline	ENSRNOG00000009514. Rattus norvegicus.
------------	--

**Family and domain databases**

InterPro	IPR000718. Peptidase_M13. IPR018497. Peptidase_M13_C. IPR008753. Peptidase_M13_N. [Graphical view]
----------	---

PANTHER	PTHR11733. Peptidase_M13. 1 hit.
---------	----------------------------------

Pfam	PF01431. Peptidase_M13. 1 hit. PF05649. Peptidase_M13_N. 1 hit. [Graphical view]
------	--

PRINTS	PR00786. NEPRILYSIN.
--------	----------------------

PROSITE	PS00142. ZINC_PROTEASE. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
NextBio	603776.
<b>Entry information</b>	
Entry name	NEP_RAT
Accession	Primary (citable) accession number: <b>P07861</b>
Entry history	Integrated into August 1, 1988 UniProtKB/Swiss-Prot: Last sequence update: January 23, 2007 Last modified: March 8, 2011 This is version 106 of the entry and version 2 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Relevant documents</b>	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact





**P19621 (NEP\_PIG) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified January 11, 2011. Version 64.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Neprilysin</b> EC=3.4.24.11 <i>Alternative name(s):</i> Atriopeptidase Enkephalinase Neutral endopeptidase 24.11 Short name=NEP Short name=Neutral endopeptidase Skin fibroblast elastase Short name=SFE CD_antigen=CD10
Gene names	Name: <b>MME</b>
Organism	<b>Sus scrofa (Pig)</b>
Taxonomic identifier	9823 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Laurasiatheria › Cetartiodactyla › Suina › Suidae › Sus

**Protein attributes**

Sequence length	26 AA.
Sequence status	Fragment.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Thermolysin-like specificity, but is almost confined on acting on polypeptides of up to 30 amino acids. Biologically important in the destruction of opioid peptides such as Met- and Leu-enkephalins by cleavage of a Gly-Phe bond. Able to cleave angiotensin-1, angiotensin-2 and angiotensin 1-9 <span>By similarity</span> . Involved in the degradation of atrial natriuretic factor (ANF). Displays UV-inducible elastase activity toward skin preelastic and elastic fibers <span>By similarity</span> .
Catalytic activity	Preferential cleavage of polypeptides between hydrophobic residues, particularly with Phe or Tyr at P1'.
Cofactor	Binds 1 zinc ion per subunit.
Subcellular location	Cell membrane; Single-pass type II membrane protein.
Sequence similarities	Belongs to the peptidase M13 family.

Ontologies

Keywords


Cellular component	Cell membrane Membrane
Domain	Transmembrane
Ligand	Zinc
Molecular function	Hydrolase Metalloprotease Protease
Technical term	Direct protein sequencing

Gene Ontology (GO)


Cellular component	integral to membrane Inferred from electronic annotation. Source: UniProtKB-KW  plasma membrane Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	metallopeptidase activity Inferred from electronic annotation. Source: UniProtKB-KW

Complete GO annotation...

Sequence annotation (Features)

Feature key	Position (s)	Length	Description	Graphical view	Feature identifier
Molecule processing					
 Chain	<1 – >26	>26	Neprilysin		PRO_000007821
Experimental info					
<input type="checkbox"/> Non-terminal residue	1	1			
<input type="checkbox"/> Non-terminal residue	26	1			

## Sequences

Sequence	Length	Mass (Da)
 P19621 [UniParc]. Last modified February 1, 1991. Version 1. Checksum: 3848804A9DDF7DEF	FASTA 26	3,039
<div><div>1020</div><div>PKPKKKQ<sup>10</sup>RWT PLEISLEVLV<sup>20</sup> LVLVXI</div></div>		
« Hide		

## References

- [1] "The N-terminal amino acid sequence of pig kidney endopeptidase-24.11 shows homology with pro-sucrase-isomaltase."  
Fulcher I.S., Pappin D.J.C., Kenny A.J.  
Biochem. J. 240:305-308(1986) [PubMed: 3548708] [Abstract]  
Cited for: PROTEIN SEQUENCE.  
Tissue: Kidney.

## Cross-references

### Sequence databases

PIR | A26070.

### 3D structure databases

ModBase | Search...

### Protein family/group databases

MEROPS | M13.001.

### Enzyme and pathway databases

BRENDA | 3.4.24.11. 249.

### Family and domain databases

PROSITE | PS00142. ZINC\_PROTEASE. Partial match.  
[Graphical view]

ProtoNet | Search...

## Entry information

Entry name	NEP_PIG
Accession	Primary (citable) accession number: <b>P19621</b>

Entry history	Integrated into February 1, 1991 UniProtKB/Swiss-Prot: Last sequence February 1, 1991 update: Last modified: January 11, 2011 This is version 64 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Relevant documents</b>	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact





**P09958 (FURIN\_HUMAN) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 143.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Furin</b> EC=3.4.21.75 <i>Alternative name(s):</i> Dibasic-processing enzyme Paired basic amino acid residue-cleaving enzyme Short name=PACE
Gene names	Name: <b>FURIN</b> Synonyms: FUR, PACE, PCSK3
Organism	<b>Homo sapiens (Human)</b> [Complete proteome]
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo

**Protein attributes**

Sequence length	794 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Furin is likely to represent the ubiquitous endoprotease activity within constitutive secretory pathways and capable of cleavage at the RX(K/R)R consensus motif. <a href="#">Ref.7</a>
Catalytic activity	Release of mature proteins from their proproteins by cleavage of -Arg-Xaa-Yaa-Arg- -Zaa- bonds, where Xaa can be any amino acid and Yaa is Arg or Lys. Releases albumin, complement component C3 and vWF from their respective precursors.
Cofactor	Calcium.
Enzyme regulation	Could be inhibited by the not secondly cleaved propeptide.
Subunit structure	Interacts with FLNA <a href="#">By similarity</a> . Binds to PACS1 which mediates TGN localization and connection to clathrin adapters. <a href="#">Ref.12</a>
Subcellular location	Golgi apparatus › trans-Golgi network membrane; Single-pass type I membrane protein. Cell membrane; Single-pass type I membrane protein. Note: Shuttles between the trans-Golgi network and the cell surface. Propeptide cleavage is a prerequisite for exit of furin molecules out of the endoplasmic

	reticulum (ER). A second cleavage within the propeptide occurs in the trans Golgi network (TGN), followed by the release of the propeptide and the activation of furin.
Tissue specificity	Seems to be expressed ubiquitously.
Domain	Contains a cytoplasmic domain responsible for its TGN localization and recycling from the cell surface. <a href="#">Ref.10</a>
Post-translational modification	<p>The inhibition peptide, which plays the role of an intramolecular chaperone, is autocatalytically removed in the endoplasmic reticulum (ER) and remains non-covalently bound to furin as a potent autoinhibitor. Following transport to the trans Golgi, a second cleavage within the inhibition propeptide results in propeptide dissociation and furin activation.</p> <p>Phosphorylation is required for TGN localization of the endoprotease. In vivo, exists as di-, mono- and non-phosphorylated forms. <a href="#">Ref.11</a></p>
Sequence similarities	<p>Belongs to the peptidase S8 family. Furin subfamily.</p> <p>Contains 1 homo B/P domain.</p>

## Ontologies

### Keywords

Cellular component	Cell membrane Golgi apparatus Membrane
Coding sequence diversity	Polymorphism
Domain	Signal Transmembrane Transmembrane helix
Ligand	Calcium Metal-binding
Molecular function	Hydrolase Protease Serine protease
PTM	Autocatalytic cleavage Cleavage on pair of basic residues Disulfide bond Glycoprotein Phosphoprotein Zymogen
Technical term	Complete proteome

### Gene Ontology (GO)


















Biological process	<p>Notch signaling pathway Traceable author statement. Source: Reactome</p> <p>cell proliferation Inferred from mutant phenotype. Source: BHF-UCL</p> <p>negative regulation of low-density lipoprotein particle receptor catabolic process Inferred from direct assay. Source: HGNC</p>
--------------------	--

	<p>negative regulation of transforming growth factor-beta1 production Inferred from mutant phenotype. Source: BHF-UCL</p> <p>nerve growth factor processing Inferred from experiment. Source: Reactome</p> <p>nerve growth factor production Inferred from direct assay. Source: BHF-UCL</p> <p>nerve growth factor receptor signaling pathway Traceable author statement. Source: Reactome</p> <p>peptide biosynthetic process Inferred from direct assay. Source: BHF-UCL</p> <p>peptidyl-glutamic acid carboxylation Traceable author statement. Source: Reactome</p> <p>positive regulation of membrane protein ectodomain proteolysis Inferred by curator. Source: BHF-UCL</p> <p>post-translational protein modification Traceable author statement. Source: Reactome</p> <p>secretion by cell Inferred from direct assay. Source: BHF-UCL</p> <p>signal peptide processing Inferred from direct assay. Source: HGNC</p> <p>transforming growth factor beta receptor signaling pathway Traceable author statement. Source: Reactome</p> <p>viral assembly, maturation, egress, and release Inferred from expression pattern. Source: BHF-UCL</p>
Cellular component	<p>Golgi lumen Inferred from experiment. Source: Reactome</p> <p>Golgi membrane Traceable author statement. Source: Reactome</p> <p>cell surface Inferred from direct assay. Source: BHF-UCL</p> <p>integral to membrane Inferred from electronic annotation. Source: UniProtKB-KW</p> <p>membrane raft Inferred from direct assay. Source: BHF-UCL</p> <p>plasma membrane Inferred from electronic annotation. Source: UniProtKB-SubCell</p> <p>trans-Golgi network Inferred from direct assay. Source: BHF-UCL</p> <p>trans-Golgi network transport vesicle Inferred from direct assay. Source: MGI</p>
Molecular function	<p>metal ion binding Inferred from electronic annotation. Source: UniProtKB-KW</p> <p>nerve growth factor binding Inferred from direct assay. Source: BHF-UCL</p> <p>peptide binding Inferred from direct assay. Source: BHF-UCL</p> <p>protease binding</p>

	Inferred from physical interaction. Source: BHF-UCL
	serine-type endopeptidase activity
	Inferred from direct assay. Source: BHF-UCL
	serine-type endopeptidase inhibitor activity
	Inferred from direct assay. Source: BHF-UCL

Complete GO annotation...

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature
Molecule processing					
	Signal peptide	1 – 24	24	Potential	
	Propeptide	25 – 107	83	Inhibition peptide	PRO_0000000000
	Chain	108 – 794	687	Furin	PRO_0000000000
Regions					
	Transmembrane	716 – 738	23	Helical; Potential	
	Region	759 – 762	4	Cell surface signal	
	Motif	498 – 500	3	Cell attachment site Potential	
	Motif	773 – 779	7	Trans Golgi network signal	
	Compositional bias	556 – 705	150	Cys-rich	
Sites					
	Active site	153	1	Charge relay system By similarity	
	Active site	194	1	Charge relay system By similarity	
	Active site	368	1	Charge relay system By similarity	
	Metal binding	115	1	Calcium 1 By similarity	
	Metal binding	162	1	Calcium 1 By similarity	
	Metal binding	208	1	Calcium 1 By similarity	
	Metal binding	258	1	Calcium 2 By similarity	
	Metal binding	301	1	Calcium 2 By similarity	
	Metal binding	331	1	Calcium 2 By similarity	



<input type="checkbox"/>	Site	75 – 76	2	Cleavage, second; by autolysis		
<input type="checkbox"/>	Site	107 – 108	2	Cleavage, first; by autolysis		
Amino acid modifications						
<input type="checkbox"/>	Modified residue	773	1	Phosphoserine; by CK2 <span>Ref.11</span>		
<input type="checkbox"/>	Modified residue	775	1	Phosphoserine; by CK2 <span>Ref.11</span>		
<input type="checkbox"/>	Glycosylation	387	1	N-linked (GlcNAc...) <span>Potential</span>		
<input type="checkbox"/>	Glycosylation	440	1	N-linked (GlcNAc...) <span>Potential</span>		
<input type="checkbox"/>	Glycosylation	553	1	N-linked (GlcNAc...) <span>Potential</span>		
<input type="checkbox"/>	Disulfide bond	211 ↔ 360		<span>By similarity</span>		
<input type="checkbox"/>	Disulfide bond	303 ↔ 333		<span>By similarity</span>		
<input type="checkbox"/>	Disulfide bond	450 ↔ 474		<span>By similarity</span>		
Natural variations						
<input type="checkbox"/>	Natural variant	43	1	A → V. <span>[dbSNP:rs16944971]</span>		VAR_()
<input type="checkbox"/>	Natural variant	547	1	W → R in cell line LoVo; does not undergo autocatalytic activation and is not transported to the Golgi apparatus. <span>Ref.8</span>		VAR_()

Sequences

Sequence	Length	Mass (Da)
<input checked="" type="checkbox"/> P09958 [UniParc]. Last modified April 1, 1990. Version 2. Checksum: 10C44DD5892EF85D	FASTA	794 86,678
<div><div><div>102030405060</div><div>MELRPWLLWVVAATGTLVLLAADAQGQKVF</div></div><div><div>708090100110120</div><div>FGDYHFWHRGVTKRSLSPHRPHSRLQRE</div></div><div><div>130140150160170180</div><div>QWYLSGVTQRDLNVKAAWAQGYTGHGIVVS</div></div><div><div>190200210220230240</div><div>DPQPRYTQMN DNRHGTRCAGEVAAVANNGV</div></div><div><div>250260270280290300</div><div>CGVGVAYNARIGGVRMLDGE</div></div><div><div></div><div>VTDAVEARSL</div></div></div>		

GLNPNHIHIY	SASWGPEDDG	KTVDGPARLA	EEAFFRGVSQ	GRGGLGSIFV	WASGNGGREH
310	320	330	340	350	360
DSCNCDGYTN	SIYTLSSISSA	TQFGNVWPYS	EACSSTLATT	YSSGNQNEKQ	IVTTDLRQKC
370	380	390	400	410	420
TESHTGTSAS	APLAAGIIAL	TLEANKNLTW	RDMQHLLVVQT	SKPAHLNAND	WATNGVGRKV
430	440	450	460	470	480
SHSYGYGLLD	AGAMVALAQN	WTTVAPQRKC	IIDILTEPKD	IGKRLEVRKT	VTACLGEPNH
490	500	510	520	530	540
ITRLEHAQAR	LTLSYNRRGD	LAIHLVSPMG	TRSTLLAARP	HDYSADGFND	WAFMTTHSWD
550	560	570	580	590	600
EDPSGEWVLE	IENTSEANNY	GTLTKFTLVL	YGTAPEGLPV	PPESSGCKTL	TSSQACVVCE
610	620	630	640	650	660
EGFSLHQKSC	VQHCPPGFAP	QVLDTHYSTE	NDVETIRASV	CAPCHASCAT	CQGPALTDCL
670	680	690	700	710	720
SCPSHASLDP	VEQTCSRQSQ	SSRESPPQQQ	PPRLPPEVEA	GQRLRAGLLP	SHLPEVVAGL
730	740	750	760	770	780
SCAFIVLVFV	TVFLVLQLRS	GFSFRGVKVY	TMDRGLISYK	GLPPEAWQEE	CPSDSEEDeg
790					
RGERTAFIKD	QSAL				

« Hide

References

	« Hide 'large scale' references
[1]	<p><b>"Structural homology between the human fur gene product and the subtilisin-like protease encoded by yeast KEX2."</b> van den Ouweland A.M.W., van Duijnhoven H.L.P., Keizer G.D., Dorssers L.C.J., van de Ven W.J.M. Nucleic Acids Res. 18:664-664(1990) [PubMed: 2408021] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA]. <u>Tissue:</u> Blood.</p>
[2]	<p><b>"Expression of a human proprotein processing enzyme: correct cleavage of the von Willebrand factor precursor at a paired basic amino acid site."</b> Wise R.J., Barr P.J., Wong P.A., Kiefer M.C., Brake A.J., Kaufman R.J. Proc. Natl. Acad. Sci. U.S.A. 87:9378-9382(1990) [PubMed: 2251280] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].</p>
[3]	<p><b>"cDNA and gene structure for a human subtilisin-like protease with cleavage specificity for paired basic amino acid residues."</b> Barr P.J., Mason O.B., Landsberg K.E., Wong P.A., Kiefer M.C., Brake A.J. DNA Cell Biol. 10:319-328(1991) [PubMed: 1713771] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].</p>
[4]	<p><b>"The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC)."</b> The MGC Project Team Genome Res. 14:2121-2127(2004) [PubMed: 15489334] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA]. <u>Tissue:</u> Lung.</p>
[5]	<p><b>"Nucleotide sequence analysis of the human fur gene."</b> Van den Ouweland A.M.W., van Groningen J.J.M., Roebroek A.J.M., Onnekink C., Van de Ven W.J.M. Nucleic Acids Res. 17:7101-7102(1989) [PubMed: 2674906] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-280.</p>

[6]	<p><b>"Evolutionary conserved close linkage of the c-fes/fps proto-oncogene and genetic sequences encoding a receptor-like protein."</b></p> <p>Roebroek A.J.M., Schalken J.A., Leunissen J.A.M., Onnekink C., Bloemers H.P.J., van de Ven W.J.M.</p> <p>EMBO J. 5:2197-2202(1986) [PubMed: 3023061] [Abstract]</p> <p><u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 296-794.</p>
[7]	<p><b>"A mutation of furin causes the lack of precursor-processing activity in human colon carcinoma LoVo cells."</b></p> <p>Takahashi S., Kasai K., Hatsuzawa K., Kitamura N., Misumi Y., Ikehara Y., Murakami K., Nakayama K.</p> <p>Biochem. Biophys. Res. Commun. 195:1019-1026(1993) [PubMed: 7690548] [Abstract]</p> <p><u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] OF 402-428, FUNCTION.</p> <p><u>Tissue:</u> Colon carcinoma.</p>
[8]	<p><b>"A second mutant allele of furin in the processing-incompetent cell line, LoVo. Evidence for involvement of the homo B domain in autocatalytic activation."</b></p> <p>Takahashi S., Nakagawa T., Kasai K., Banno T., Duguay S.J., Van de Ven W.J.M., Murakami K., Nakayama K.</p> <p>J. Biol. Chem. 270:26565-26569(1995) [PubMed: 7592877] [Abstract]</p> <p><u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] OF 527-553, VARIANT ARG-547.</p> <p><u>Tissue:</u> Colon carcinoma.</p>
[9]	<p><b>"Activation of human furin precursor processing endoprotease occurs by an intramolecular autoproteolytic cleavage."</b></p> <p>Leduc R., Molloy S.S., Thorne B.A., Thomas G.</p> <p>J. Biol. Chem. 267:14304-14308(1992) [PubMed: 1629222] [Abstract]</p> <p><u>Cited for:</u> PROTEOLYTIC PROCESSING.</p>
[10]	<p><b>"Homology modelling of the catalytic domain of human furin. A model for the eukaryotic subtilisin-like proprotein convertases."</b></p> <p>Siezen R.J., Creemers J.W.M., van de Ven W.J.M.</p> <p>Eur. J. Biochem. 222:255-266(1994) [PubMed: 8020465] [Abstract]</p> <p><u>Cited for:</u> 3D-STRUCTURE MODELING OF CATALYTIC DOMAIN.</p>
[11]	<p><b>"Intracellular trafficking of furin is modulated by the phosphorylation state of a casein kinase II site in its cytoplasmic tail."</b></p> <p>Jones B.G., Thomas L., Molloy S.S., Thulin C.D., Fry M.D., Walsh K.A., Thomas G.</p> <p>EMBO J. 14:5869-5883(1995) [PubMed: 8846780] [Abstract]</p> <p><u>Cited for:</u> PHOSPHORYLATION AT SER-773 AND SER-775.</p>
[12]	<p><b>"PACS-1 binding to adaptors is required for acidic cluster motif-mediated protein traffic."</b></p> <p>Crump C.M., Xiang Y., Thomas L., Gu F., Austin C., Tooze S.A., Thomas G.</p> <p>EMBO J. 20:2191-2201(2001) [PubMed: 11331585] [Abstract]</p> <p><u>Cited for:</u> INTERACTION WITH PACS1.</p>
+	Additional computationally mapped references.

## Cross-references

### Sequence databases

<ul style="list-style-type: none"> <li>⊙ EMBL</li> <li>⊙ GenBank</li> <li>⊙ DDBJ</li> </ul>	<p>X17094 mRNA. Translation: CAA34948.1.</p> <p>BC012181 mRNA. Translation: AAH12181.1.</p> <p>X15723 Genomic DNA. Translation: CAA33745.1.</p> <p>X04329 Genomic DNA. Translation: CAA27860.1.</p>
IPI	IPI00018387.
PIR	KXHUF. A39552.

RefSeq	NP_002560.1. NM_002569.2.
UniGene	Hs.513153.
<b>3D structure databases</b>	
ProteinModelPortal	P09958.
SMR	P09958. Positions 30-100, 109-578, 592-689.
ModBase	Search...
<b>Protein-protein interaction databases</b>	
DIP	DIP-29904N.
IntAct	P09958. 1 interaction.
MINT	MINT-1209355.
STRING	P09958.
<b>Protein family/group databases</b>	
MEROPS	S08.071.
<b>PTM databases</b>	
PhosphoSite	P09958.
<b>2-D gel databases</b>	
OGP	P09958.
<b>Proteomic databases</b>	
PeptideAtlas	P09958.
PRIDE	P09958.
<b>Genome annotation databases</b>	
Ensembl	ENST00000268171; ENSP00000268171; ENSG00000140564.
GeneID	5045.
KEGG	hsa:5045.
UCSC	uc002bpu.1. human.
<b>Organism-specific databases</b>	
CTD	5045.
GeneCards	GC15P067523.
H-InvDB	HIX0202168.
HGNC	HGNC:8568. FURIN.
HPA	CAB009499.
MIM	136950. gene.
neXtProt	NX_P09958.
PharmGKB	PA32894.



GenAtlas	Search...
<b>Phylogenomic databases</b>	
eggNOG	prNOG13720.
GeneTree	ENSGT00600000084064.
HOGENOM	HBG715943.
HOVERGEN	HBG008705.
InParanoid	P09958.
OMA	KHGFLNL.
OrthoDB	EOG4ZW59M.
PhylomeDB	P09958.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.21.75. 247.
Pathway_Interaction_DB	glypican_3pathway. Glypican 3 network. hif1_tfpathway. HIF-1-alpha transcription factor network. p75ntrpathway. p75(NTR)-mediated signaling.
Reactome	REACT_11061. Signalling by NGF. REACT_16888. Signaling by PDGF. REACT_17015. Metabolism of proteins. REACT_299. Signaling by Notch. REACT_6844. Signaling by TGF beta.
<b>Gene expression databases</b>	
ArrayExpress	P09958.
Bgee	P09958.
CleanEx	HS_FURIN.
Genevestigator	P09958.
GermOnline	ENSG00000140564. Homo sapiens.
<b>Family and domain databases</b>	
InterPro	IPR006212. Furin_repeat. IPR008979. Galactose-bd-like. IPR009030. Growth_fac_rcpt. IPR000209. Peptidase_S8/S53. IPR022398. Peptidase_S8/S53_AS. IPR015500. Peptidase_S8_subtilisin-rel. IPR009020. Prot_inh_propept. IPR002884. PrprotnconvertsP. [Graphical view]
Gene3D	G3DSA:3.40.50.200. Pept_S8_S53. 1 hit.
PANTHER	PTHR10795. SubtilSerProt. 1 hit.
Pfam	PF01483. P_proprotein. 1 hit. PF00082. Peptidase_S8. 1 hit. [Graphical view]
PRINTS	PR00723. SUBTILISIN.
SMART	SM00261. FU. 2 hits. [Graphical view]

SUPFAM	SSF49785. Gal_bind_like. 1 hit. SSF57184. Grow_fac_recept. 1 hit. SSF52743. Pept_S8_S53. 1 hit. SSF54897. Prot_inh_propept. 1 hit.
PROSITE	PS00136. SUBTILASE_ASP. 1 hit. PS00137. SUBTILASE_HIS. 1 hit. PS00138. SUBTILASE_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
NextBio	19422.
PMAP-CutDB	P09958.
SOURCE	Search...
<b>Entry information</b>	
Entry name	FURIN_HUMAN
Accession	Primary (citable) accession number: <b>P09958</b> Secondary accession number(s): Q14336, Q6LBS3, Q9UCZ5
Entry history	Integrated into July 1, 1989 UniProtKB/Swiss-Prot: Last sequence April 1, 1990 update: Last modified: April 5, 2011 This is version 143 of the entry and version 2 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
Disclaimer	Any medical or genetic information present in this entry is provided for research, educational and informational purposes only. It is not in any way intended to be used as a substitute for professional medical advice, diagnosis, treatment or care.
<b>Relevant documents</b>	
Human chromosome 15 Human chromosome 15: entries, gene names and cross-references to MIM	
Human entries with polymorphisms or disease mutations List of human entries with polymorphisms or disease mutations	
Human polymorphisms and disease mutations Index of human polymorphisms and disease mutations	
MIM cross-references Online Mendelian Inheritance in Man (MIM) cross-references in UniProtKB/Swiss-Prot	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



**P23188 (FURIN\_MOUSE) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified March 8, 2011. Version 114.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Furin</b> EC=3.4.21.75 <i>Alternative name(s):</i> Dibasic-processing enzyme Paired basic amino acid residue-cleaving enzyme Short name=PACE Prohormone convertase 3
Gene names	Name: <b>Furin</b> Synonyms: Fur, Pcsk3
Organism	<b>Mus musculus (Mouse)</b>
Taxonomic identifier	10090 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires › Rodentia › Sciurognathi › Muroidea › Muridae › Murinae › Mus › Mus

**Protein attributes**

Sequence length	793 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Furin is likely to represent the ubiquitous endoprotease activity within constitutive secretory pathways and capable of cleavage at the RX(K/R)R consensus motif.
Catalytic activity	Release of mature proteins from their proproteins by cleavage of -Arg-Xaa-Yaa-Arg- -Zaa- bonds, where Xaa can be any amino acid and Yaa is Arg or Lys. Releases albumin, complement component C3 and vWF from their respective precursors.
Cofactor	Binds 2 calcium ions per subunit.
Enzyme regulation	Could be inhibited by the not secondly cleaved propeptide.
Subunit structure	Interacts with FLNA <span>By similarity</span> . Binds to PACS1 which mediates TGN localization and connection to clathrin adaptors <span>By similarity</span> <span>Ref.3</span>
Subcellular location	Golgi apparatus › trans-Golgi network membrane; Single-pass type I membrane protein. Cell membrane; Single-pass type I membrane protein. Note: Shuttles between the trans-Golgi



	network and the cell surface. Propeptide cleavage is a prerequisite for exit of furin molecules out of the endoplasmic reticulum (ER). A second cleavage within the propeptide occurs in the trans Golgi network (TGN), followed by the release of the propeptide and the activation of furin <span>By similarity</span> .
Tissue specificity	Seems to be expressed ubiquitously.
Domain	Contains a cytoplasmic domain responsible for its TGN localization and recycling from the cell surface.
Post-translational modification	<p>The inhibition peptide, which plays the role of an intramolecular chaperone, is autocatalytically removed in the endoplasmic reticulum (ER) and remains non-covalently bound to furin as a potent autoinhibitor. Following transport to the trans Golgi, a second cleavage within the inhibition propeptide results in propeptide dissociation and furin activation <span>By similarity</span>.</p> <p>Phosphorylation is required for TGN localization of the endoprotease. In vivo, exists as di-, mono- and non-phosphorylated forms <span>By similarity</span>.</p>
Sequence similarities	<p>Belongs to the peptidase S8 family. Furin subfamily.</p> <p>Contains 1 homo B/P domain.</p>
<b>Ontologies</b>	
<b>Keywords</b>	
Cellular component	Cell membrane Golgi apparatus Membrane
Domain	Signal Transmembrane Transmembrane helix
Ligand	Calcium Metal-binding
Molecular function	Hydrolase Protease Serine protease
PTM	Autocatalytic cleavage Cleavage on pair of basic residues Disulfide bond Glycoprotein Phosphoprotein Zymogen
Technical term	3D-structure
<b>Gene Ontology (GO)</b>	
Cellular component	<p>early endosome Traceable author statement. Source: MGI</p> <p>endoplasmic reticulum lumen Traceable author statement. Source: Reactome</p> <p>integral to membrane Traceable author statement. Source: MGI</p>

## Molecular function

plasma membrane  
Inferred from electronic annotation. Source: UniProtKB-SubCell

trans-Golgi network transport vesicle membrane  
Traceable author statement. Source: MGI

---

metal ion binding  
Inferred from electronic annotation. Source: UniProtKB-KW

Complete GO annotation...

## Sequence annotation (Features)

	Feature key	Position(s)	Length	Description	Graphical view	Feature id
<b>Molecule processing</b>						
<input checked="" type="checkbox"/>	Signal peptide	1 – 24	24	Potential		
<input checked="" type="checkbox"/>	Propeptide	25 – 107	83	Inhibition peptide By similarity		PRO_000001
<input checked="" type="checkbox"/>	Chain	108 – 793	686	Furin		PRO_000001
<b>Regions</b>						
<input checked="" type="checkbox"/>	Transmembrane	715 – 735	21	Helical; Potential		
<input checked="" type="checkbox"/>	Region	109 – 445	337	Catalytic		
<input checked="" type="checkbox"/>	Region	446 – 578	133	P-domain		
<input type="checkbox"/>	Region	758 – 761	4	Cell surface signal		
<input type="checkbox"/>	Motif	498 – 500	3	Cell attachment site Potential		
<input type="checkbox"/>	Motif	772 – 778	7	Trans Golgi network signal		
<input checked="" type="checkbox"/>	Compositional bias	556 – 705	150	Cys-rich		
<b>Sites</b>						
<input type="checkbox"/>	Active site	153	1	Charge relay system		
<input type="checkbox"/>	Active site	194	1	Charge relay system		
<input type="checkbox"/>	Active site	368	1	Charge relay system		
<input type="checkbox"/>	Metal binding	115	1	Calcium 1		
<input type="checkbox"/>	Metal binding	162	1	Calcium 1		
<input type="checkbox"/>	Metal binding	208	1	Calcium 1		
<input type="checkbox"/>	Metal binding	258	1	Calcium 2		

<input type="checkbox"/>	Metal binding	301	1	Calcium 2		
<input type="checkbox"/>	Metal binding	331	1	Calcium 2		
<input type="checkbox"/>	Site	75 – 76	2	Cleavage, second; by autolysis <div>By similarity</div>		
<input type="checkbox"/>	Site	107 – 108	2	Cleavage, first; by autolysis <div>By similarity</div>		

Amino acid modifications

<input type="checkbox"/>	Modified residue	772	1	Phosphoserine; by CK2 <div>By similarity</div>		
<input type="checkbox"/>	Modified residue	774	1	Phosphoserine; by CK2 <div>By similarity</div>		
<input type="checkbox"/>	Glycosylation	387	1	N-linked (GlcNAc...)		
<input type="checkbox"/>	Glycosylation	440	1	N-linked (GlcNAc...)		
<input type="checkbox"/>	Glycosylation	553	1	N-linked (GlcNAc...) <div>Potential</div>		
<input type="checkbox"/>	Disulfide bond	211 ↔ 360				
<input type="checkbox"/>	Disulfide bond	303 ↔ 333				
<input type="checkbox"/>	Disulfide bond	450 ↔ 474				

Experimental info

<input type="checkbox"/>	Sequence conflict	746	1	M → V in AAA37643. <div>Ref.2</div>		
--------------------------	-------------------	-----	---	--	--	--

Secondary structure

1 . .....  

■

 Helix 

■

 Strand 

■

 Turn

Details...

Sequences

Sequence	Length	Mass (Da)
<input type="checkbox"/> P23188 [UniParc]. Last modified November 1, 1991. Version 1. Checksum: 5F121C3DE2E1A42D	FASTA 793	86,804

```

      10      20      30      40      50      60
MELRSWLLWV VAAAGAVVLL AADAQGQKIF TNTWAVHIPG GPAVADRVAQ KHGFHNLGQI

      70      80      90     100     110     120
FGDYHFWHR AVTKRSLSPH RPRHSRLQRE PQVKWLEQQV AKRRAKRDVY QEPTDPKFPQ

     130     140     150     160     170     180
QWYLSGVTQR DLNVKEAWAQ GFTGHGIVVS ILDDGIEKNH PDLAGNYDPG ASFDVNDQDP

     190     200     210     220     230     240
DPQPRYTQMN DNRHGTRCAG EVAAVANNGV CGVGVAYNAR IGGVRMLDGE VTDAVEARSL

     250     260     270     280     290     300
GLNPNHIHIY SASWGPEDDG KTVDGPARLA EEAFFRGVSQ GRGGLGSIFV WASGNGGREH

     310     320     330     340     350     360
DSCNCDGYTN SIYTLSSISA TQFGNVPWYS EACSSTLATT YSSGNQNEKQ IVTTDLRQKC

     370     380     390     400     410     420
TESHTGTSAS APLAAGIIAL TLEANKNLTW RDMQHLVVQT SKPAHLNADD WATNGVGRKV

     430     440     450     460     470     480
SHSYGYGLLD AGAMVALAQN WTTVAPQRKC IVEILVEPKD IGRLEVRKA VTACLGEPNH

     490     500     510     520     530     540
ITRLEHVQAR LTLSYNRRGD LAIHLISPMG TRSTLLAARP HDYSADGFND WAFMTTHSWD

     550     560     570     580     590     600
EDPAGEWVLE IENTSEANNY GTLTKFTLVL YGTAPEGLST PPSSGCKTL TSSQACVVCE

     610     620     630     640     650     660
EGYSLHQKSC VQHCPPGFIP QVLDTHYSTE NDVEIIRASV CTPCHASCAT CQGPAPTDCL

     670     680     690     700     710     720
SCPSHASLDP VEQTCSRQSQ SSRESRPQQQ PPALRPEVEM EPRLQAGLAS HLPEVLAGLS

     730     740     750     760     770     780
CLIIVLIFGI VFLFLHRC SG FSFRGMKVYT MDRGLISYKG LPPEAWQEEC PSDSEDEGR

     790
GERTAFIKDQ SAL

```

« Hide

## References

- [1] **"Structure and expression of mouse furin, a yeast Kex2-related protease. Lack of processing of coexpressed prorenin in GH4C1 cells."**  
Hatsuzawa K., Hosaka M., Nakagawa T., Nagase M., Shoda A., Murakami K., Nakayama K.  
J. Biol. Chem. 265:22075-22078(1990) [PubMed: 2266110] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA].
- [2] **"Cloning and functional expression of a 4.3 kbp mouse fur cDNA: evidence for differential expression."**  
Creemers J.W.M., Roebroek A.J.M., van den Ouweland A.M.W., van Duijnhoven H.L.P., van de Ven W.J.M.  
Life Sci. Adv. (Mol. Biol.) 11:127-138(1992)  
Cited for: NUCLEOTIDE SEQUENCE [MRNA].  
Tissue: Liver.
- [3] **"Cytoskeletal protein ABP-280 directs the intracellular trafficking of furin and modulates proprotein processing in the endocytic pathway."**  
Liu G., Thomas L., Warren R.A., Enns C.A., Cunningham C.C., Hartwig J.H., Thomas G.  
J. Cell Biol. 139:1719-1733(1997) [PubMed: 9412467] [Abstract]  
Cited for: INTERACTION WITH FLNA.
- [4] **"The crystal structure of the proprotein processing proteinase furin explains its stringent specificity."**  
Henrich S., Cameron A., Bourenkov G.P., Kiefersauer R., Huber R., Lindberg I., Bode W., Than M.E.



Nat. Struct. Biol. 10:520-526(2003) [PubMed: 12794637] [Abstract]  
Cited for: X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 108-578 IN COMPLEX WITH INHIBITOR.

+ Additional computationally mapped references.

## Cross-references

### Sequence databases

<input checked="" type="radio"/> EMBL <input checked="" type="radio"/> GenBank <input checked="" type="radio"/> DDBJ	X54056 mRNA. Translation: CAA37988.1. L26489 mRNA. Translation: AAA37643.1.
--	--

IPI	IPI00314624.
-----	--------------

PIR	KXMSF. A23679.
-----	----------------

RefSeq	NP_001074923.1. NM_001081454.1. NP_035176.1. NM_011046.2.
--------	--

UniGene	Mm.5241.
---------	----------

### 3D structure databases

<input checked="" type="radio"/> PDBe <input checked="" type="radio"/> RCSB PDB <input checked="" type="radio"/> PDBj	<table border="0"> <tr> <th>Entry</th> <th>Method</th> <th>Resolution (Å)</th> <th>Chain</th> <th>Positions</th> <th>PDBsum</th> </tr> <tr> <td>1P8J</td> <td>X-ray</td> <td>2.60</td> <td>A/B/C/D/E/F/G/H</td> <td>108-578</td> <td>[&gt;]</td> </tr> </table>	Entry	Method	Resolution (Å)	Chain	Positions	PDBsum	1P8J	X-ray	2.60	A/B/C/D/E/F/G/H	108-578	[>]
Entry	Method	Resolution (Å)	Chain	Positions	PDBsum								
1P8J	X-ray	2.60	A/B/C/D/E/F/G/H	108-578	[>]								

ProteinModelPortal	P23188.
--------------------	---------

SMR	P23188. Positions 30-100, 109-578, 593-681.
-----	---

ModBase	Search...
---------	-----------

### Protein-protein interaction databases

STRING	P23188.
--------	---------

### Protein family/group databases

MEROPS	S08.071.
--------	----------

### PTM databases

PhosphoSite	P23188.
-------------	---------

### Proteomic databases

PRIDE	P23188.
-------	---------

### Genome annotation databases

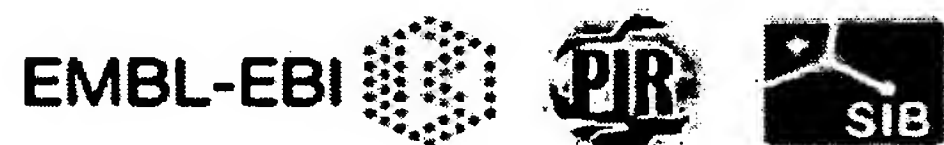
Ensembl	ENSMUST00000107362; ENSMUSP00000102985; ENSMUSG00000030530. ENSMUST00000120753; ENSMUSP00000113793; ENSMUSG00000030530. ENSMUST00000122232; ENSMUSP00000113370; ENSMUSG00000030530.
---------	--

GeneID	18550.
--------	--------

KEGG	mmu:18550.
<b>Organism-specific databases</b>	
CTD	18550.
MGI	MGI:97513. Furin.
<b>Phylogenomic databases</b>	
eggNOG	roNOG10014.
HOGENOM	HBG715943.
HOVERGEN	HBG008705.
InParanoid	P23188.
OrthoDB	EOG4ZW59M.
PhylomeDB	P23188.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.21.75. 244.
<b>Gene expression databases</b>	
ArrayExpress	P23188.
Bgee	P23188.
CleanEx	MM_FURIN.
Genevestigator	P23188.
GermOnline	ENSMUSG00000030530. <i>Mus musculus</i> .
<b>Family and domain databases</b>	
InterPro	IPR006212. Furin_repeat. IPR008979. Galactose-bd-like. IPR009030. Growth_fac_rcpt. IPR000209. Peptidase_S8/S53. IPR022398. Peptidase_S8/S53_AS. IPR015500. Peptidase_S8_subtilisin-rel. IPR009020. Prot_inh_propept. IPR002884. PrprotnconvertsP. [Graphical view]
Gene3D	G3DSA:3.40.50.200. Pept_S8_S53. 1 hit.
PANTHER	PTHR10795. SubtilSerProt. 1 hit.
Pfam	PF01483. P_proprotein. 1 hit. PF00082. Peptidase_S8. 1 hit. [Graphical view]
PRINTS	PR00723. SUBTILISIN.
SMART	SM00261. FU. 2 hits. [Graphical view]
SUPFAM	SSF49785. Gal_bind_like. 1 hit. SSF57184. Grow_fac_recept. 1 hit. SSF52743. Pept_S8_S53. 1 hit. SSF54897. Prot_inh_propept. 1 hit.
PROSITE	PS00136. SUBTILASE_ASP. 1 hit. PS00137. SUBTILASE_HIS. 1 hit.

	PS00138. SUBTILASE_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
SOURCE	Search...
<b>Entry information</b>	
Entry name	FURIN_MOUSE
Accession	Primary (citable) accession number: <b>P23188</b>
Entry history	Integrated into November 1, 1991 UniProtKB/Swiss-Prot: Last sequence November 1, 1991 update: Last modified: March 8, 2011 This is version 114 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Relevant documents</b>	
MGD cross-references Mouse Genome Database (MGD) cross-references in UniProtKB/Swiss-Prot	
PDB cross-references Index of Protein Data Bank (PDB) cross-references	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License &amp; Disclaimer | Contact



**Q28193 (FURIN\_BOVIN) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 100.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Furin</b> EC=3.4.21.75 <i>Alternative name(s):</i> Dibasic-processing enzyme Paired basic amino acid residue-cleaving enzyme Short name=PACE Trans Golgi network protease furin
Gene names	Name: <b>FURIN</b> Synonyms: FUR, PACE
Organism	<b>Bos taurus (Bovine)</b>
Taxonomic identifier	9913 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Laurasiatheria › Cetartiodactyla › Ruminantia › Pecora › Bovidae › Bovinae › Bos

**Protein attributes**

Sequence length	797 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Furin is likely to represent the ubiquitous endoprotease activity within constitutive secretory pathways and capable of cleavage at the RX(K/R)R consensus motif.
Catalytic activity	Release of mature proteins from their proproteins by cleavage of -Arg-Xaa-Yaa-Arg- -Zaa- bonds, where Xaa can be any amino acid and Yaa is Arg or Lys. Releases albumin, complement component C3 and vWF from their respective precursors.
Cofactor	Calcium <span>By similarity</span> .
Enzyme regulation	Could be inhibited by the not secondly cleaved propeptide.
Subunit structure	Interacts with FLNA. Binds to PACS1 which mediates TGN localization and connection to clathrin adapters <span>By similarity</span> .
Subcellular location	Golgi apparatus › trans-Golgi network membrane; Single-pass type I membrane protein. Cell membrane; Single-pass type I membrane protein. Note: Shuttles between the trans-Golgi network and the cell surface. Propeptide cleavage is a



	<p>prerequisite for exit of furin molecules out of the endoplasmic reticulum (ER). A second cleavage within the propeptide occurs in the trans Golgi network (TGN), followed by the release of the propeptide and the activation of furin.</p>
Tissue specificity	<p>Seems to be expressed ubiquitously.</p>
Domain	<p>Contains a cytoplasmic domain responsible for its TGN localization and recycling from the cell surface.</p>
Post-translational modification	<p>The inhibition peptide, which plays the role of an intramolecular chaperone, is autocatalytically removed in the endoplasmic reticulum (ER) and remains non-covalently bound to furin as a potent autoinhibitor. Following transport to the trans Golgi, a second cleavage within the inhibition propeptide results in propeptide dissociation and furin activation.</p> <p>Phosphorylation is required for TGN localization of the endoprotease. In vivo, exists as di-, mono- and non-phosphorylated forms <span>By similarity</span>.</p>
Sequence similarities	<p>Belongs to the peptidase S8 family. Furin subfamily.</p> <p>Contains 1 homo B/P domain.</p>

Ontologies

Keywords





Cellular component	Cell membrane Golgi apparatus Membrane
Domain	Signal Transmembrane Transmembrane helix
Ligand	Calcium Metal-binding
Molecular function	Hydrolase Protease Serine protease
PTM	Autocatalytic cleavage Cleavage on pair of basic residues Disulfide bond Glycoprotein Phosphoprotein Zymogen

Gene Ontology (GO)

Cellular component	integral to membrane Inferred from electronic annotation. Source: UniProtKB-KW plasma membrane Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	metal ion binding Inferred from electronic annotation. Source: UniProtKB-KW

Complete GO annotation...

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature id
Molecule processing					
 Signal peptide	1 – 24	24	Potential		
 Propeptide	25 – 107	83	Inhibition peptide By similarity		PRO_000001
 Chain	108 – 797	690	Furin		PRO_000001
Regions					
 Transmembrane	719 – 741	23	Helical; Potential		

<input type="checkbox"/>	Region	762 – 765	4	Cell surface signal		
<input type="checkbox"/>	Motif	498 – 500	3	Cell attachment site <small>Potential</small>		
<input type="checkbox"/>	Motif	776 – 782	7	Trans Golgi network signal		
<input checked="" type="checkbox"/>	Compositional bias	556 – 708	153	Cys-rich		
<input type="checkbox"/>	Compositional bias	690 – 698	9	Poly-Pro		
<b>Sites</b>						
<input type="checkbox"/>	Active site	153	1	Charge relay system <small>By similarity</small>		
<input type="checkbox"/>	Active site	194	1	Charge relay system <small>By similarity</small>		
<input type="checkbox"/>	Active site	368	1	Charge relay system <small>By similarity</small>		
<input type="checkbox"/>	Metal binding	115	1	Calcium 1 <small>By similarity</small>		
<input type="checkbox"/>	Metal binding	162	1	Calcium 1 <small>By similarity</small>		
<input type="checkbox"/>	Metal binding	208	1	Calcium 1 <small>By similarity</small>		
<input type="checkbox"/>	Metal binding	258	1	Calcium 2 <small>By similarity</small>		
<input type="checkbox"/>	Metal binding	301	1	Calcium 2 <small>By similarity</small>		
<input type="checkbox"/>	Metal binding	331	1	Calcium 2 <small>By similarity</small>		
<input type="checkbox"/>	Site	75 – 76	2	Cleavage, second; by autolysis <small>By similarity</small>		
<input type="checkbox"/>	Site	107 – 108	2	Cleavage, first; by autolysis <small>By similarity</small>		
<b>Amino acid modifications</b>						
<input type="checkbox"/>	Modified residue	776	1	Phosphoserine; by CK2 <small>By similarity</small>		
<input type="checkbox"/>	Modified residue	778	1	Phosphoserine; by CK2 <small>By similarity</small>		

<input type="checkbox"/>	Glycosylation	387	1	N-linked (GlcNAc...) Potential	I	
<input type="checkbox"/>	Glycosylation	440	1	N-linked (GlcNAc...) Potential	I	
<input type="checkbox"/>	Glycosylation	553	1	N-linked (GlcNAc...) Potential	I	
<input type="checkbox"/>	Disulfide bond	211 ↔ 360		By similarity	I I	
<input type="checkbox"/>	Disulfide bond	303 ↔ 333		By similarity	II	
<input type="checkbox"/>	Disulfide bond	450 ↔ 474		By similarity	II	

Sequences

Sequence	Length	Mass (Da)
<div><div><input type="checkbox"/></div><div>Q28193 [UniParc]. Last modified November 1, 1997. Version 1. Checksum: 466F28EC0246C3D2</div></div>	FASTA 797	87,251
<div><div><div><div>10</div><div>20</div><div>30</div><div>40</div><div>50</div><div>60</div></div><div>MELRPWLFV VAAAGALVLL VADARGEKVF TNTWAVHIPG GPAVADRVAR KHGFLNLGQI</div></div><div><div><div>70</div><div>80</div><div>90</div><div>100</div><div>110</div><div>120</div></div><div>FGDYHFWHR AVTKRSLSPH RLGHNRLORE PQVKWLEQQV AKRRAKRDYI QEPTDPKFPQ</div></div><div><div><div>130</div><div>140</div><div>150</div><div>160</div><div>170</div><div>180</div></div><div>QWYLSGVTQR DLNVKEAWAQ GYTGRGIVVS ILDDGIEKNH PDLAGNYDPG ASFDVNDQDP</div></div><div><div><div>190</div><div>200</div><div>210</div><div>220</div><div>230</div><div>240</div></div><div>DPQPRYTQMN DNRHGTRCAG EVAAVANNGV CGVGVAYNAR IGGVRMLDGE VTDAVEARSL</div></div><div><div><div>250</div><div>260</div><div>270</div><div>280</div><div>290</div><div>300</div></div><div>GLNPNHIHIY SASWGPEDDG KTVDPGPAHLA EEAFRQVSVQ GRGGLGSIFV WASGNGGREH</div></div><div><div><div>310</div><div>320</div><div>330</div><div>340</div><div>350</div><div>360</div></div><div>DSCNCDGYTN SIYTLSSISA TQFGNVPWYS EACSSTLATY YSSGNQNEKQ IVTTDLRQKC</div></div><div><div><div>370</div><div>380</div><div>390</div><div>400</div><div>410</div><div>420</div></div><div>TESHTGTSAF APLAAGIIAL TLEANKNLTW RDMQHLVVRT SKPAHLNAND WATNGVGRKV</div></div><div><div><div>430</div><div>440</div><div>450</div><div>460</div><div>470</div><div>480</div></div><div>SHSYGYGLLD AGAMVALAQN WTTVAPQRKC TIDILTEPKD IGRLEVRKT VTACLGEPSH</div></div><div><div><div>490</div><div>500</div><div>510</div><div>520</div><div>530</div><div>540</div></div><div>ITRLEHAQAR LTLSYNRRGD LAIHLVSPMG TRSTLLAARP HDYSADGFND WAFMTTHSWD</div></div><div><div><div>550</div><div>560</div><div>570</div><div>580</div><div>590</div><div>600</div></div><div>EDPSGEWVLE IENTSEANNY GTLTKEFTLV YGTAPEGLPT PPESIGCKTL TSSQACVVCE</div></div><div><div><div>610</div><div>620</div><div>630</div><div>640</div><div>650</div><div>660</div></div><div>EGFSLHQKNC VQHCPPGFAP QVLDTHYSTE NDVEIIRASV CTPCHASCAT CQGPAPTDCL</div></div><div><div><div>670</div><div>680</div><div>690</div><div>700</div><div>710</div><div>720</div></div><div>SCPSHASLDP VEQTCSRQSQ SSRESHQQQP PPPPRPPPAE VATEPRLRAD LLPSHLPEVV</div></div><div><div><div>730</div><div>740</div><div>750</div><div>760</div><div>770</div><div>780</div></div><div>AGLSCAFIVL VFVTVFLVLQ LRSGFSFRGV KVTMDRGLI SYKGLPPEAW QEECPDSEE</div></div><div><div><div>790</div></div><div>DEGRGERTAF IKDQSAL</div></div></div>		

« Hide



## References

- [1] "Maturation of the trans-Golgi network protease furin: compartmentalization of propeptide removal, substrate cleavage, and COOH-terminal truncation."  
 Vey M., Schaefer W., Berghoefer S., Klenk H., Garten W.  
 J. Cell Biol. 127:1829-1842(1994) [PubMed: 7806563] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA], PROTEOLYTIC PROCESSING.  
Tissue: Kidney.

## Cross-references

### Sequence databases

<input type="radio"/> EMBL <input type="radio"/> GenBank <input type="radio"/> DDBJ	X75956 mRNA. Translation: CAA53569.1.
---	---------------------------------------

IPI	IPI00685701.
PIR	I46044.
RefSeq	NP_776561.1. NM_174136.2.
UniGene	Bt.21338.

### 3D structure databases

ProteinModelPortal	Q28193.
SMR	Q28193. Positions 30-99, 110-575, 592-665.
ModBase	Search...

### Protein-protein interaction databases

STRING	Q28193.
--------	---------

### Protein family/group databases

MEROPS	S08.071.
--------	----------

### Proteomic databases

PRIDE	Q28193.
-------	---------

### Genome annotation databases

Ensembl	ENSBTAT00000003826; ENSBTAP00000003826; ENSBTAG00000002939.
GeneID	281374.
KEGG	bta:281374.

### Organism-specific databases

CTD	281374.
-----	---------

<b>Phylogenomic databases</b>	
eggNOG	maNOG16185.
GeneTree	ENSGT00600000084064.
HOVERGEN	HBG008705.
InParanoid	Q28193.
OrthoDB	EOG4ZW59M.
PhylomeDB	Q28193.
<b>Enzyme and pathway databases</b>	
BioCyc	CATTLE:281374-MONOMER.
BRENDA	3.4.21.75. 251.
<b>Family and domain databases</b>	
InterPro	IPR006212. Furin_repeat. IPR008979. Galactose-bd-like. IPR009030. Growth_fac_rcpt. IPR000209. Peptidase_S8/S53. IPR022398. Peptidase_S8/S53_AS. IPR015500. Peptidase_S8_subtilisin-rel. IPR009020. Prot_inh_propept. IPR002884. PrprotncnvertsP. [Graphical view]
Gene3D	G3DSA:3.40.50.200. Pept_S8_S53. 1 hit.
PANTHER	PTHR10795. SubtilSerProt. 1 hit.
Pfam	PF01483. P_proprotein. 1 hit. PF00082. Peptidase_S8. 1 hit. [Graphical view]
PRINTS	PR00723. SUBTILISIN.
SMART	SM00261. FU. 2 hits. [Graphical view]
SUPFAM	SSF49785. Gal_bind_like. 1 hit. SSF57184. Grow_fac_recept. 1 hit. SSF52743. Pept_S8_S53. 1 hit. SSF54897. Prot_inh_propept. 1 hit.
PROSITE	PS00136. SUBTILASE_ASP. 1 hit. PS00137. SUBTILASE_HIS. 1 hit. PS00138. SUBTILASE_SER. False negative. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
PMAP-CutDB	Q28193.
<b>Entry information</b>	
Entry name	FURIN_BOVIN
Accession	Primary (citable) accession number: <b>Q28193</b>

Entry history	Integrated into November 1, 1997 UniProtKB/Swiss-Prot: Last sequence update: November 1, 1997 Last modified: April 5, 2011 This is version 100 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Relevant documents</b>	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



**P23377 (FURIN\_RAT) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified March 8, 2011. Version 110.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Furin</b> EC=3.4.21.75 <i>Alternative name(s):</i> Dibasic-processing enzyme Paired basic amino acid residue-cleaving enzyme Short name=PACE Prohormone convertase 3
Gene names	Name: <b>Furin</b> Synonyms: Fur, Pcsk3
Organism	<b>Rattus norvegicus (Rat)</b>
Taxonomic identifier	10116 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires › Rodentia › Sciurognathi › Muroidea › Muridae › Murinae › Rattus

**Protein attributes**

Sequence length	793 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Furin is likely to represent the ubiquitous endoprotease activity within constitutive secretory pathways and capable of cleavage at the RX(K/R)R consensus motif.
Catalytic activity	Release of mature proteins from their proproteins by cleavage of -Arg-Xaa-Yaa-Arg- -Zaa- bonds, where Xaa can be any amino acid and Yaa is Arg or Lys. Releases albumin, complement component C3 and vWF from their respective precursors.
Cofactor	Calcium <span>By similarity</span> .
Enzyme regulation	Could be inhibited by the not secondly cleaved propeptide.
Subunit structure	Interacts with FLNA <span>By similarity</span> . Binds to PACS1 which mediates TGN localization and connection to clathrin adapters. <span>Ref.2</span>
Subcellular location	Golgi apparatus › trans-Golgi network membrane; Single-pass type I membrane protein. Cell membrane; Single-pass type I membrane protein. Note: Shuttles between the trans-Golgi



	network and the cell surface. Propeptide cleavage is a prerequisite for exit of furin molecules out of the endoplasmic reticulum (ER). A second cleavage within the propeptide occurs in the trans Golgi network (TGN), followed by the release of the propeptide and the activation of furin <span>By similarity</span> .
Tissue specificity	Seems to be expressed ubiquitously.
Developmental stage	Expressed at E7 day in endoderm and mesoderm, uniformly expressed until E10, when expression is higher in heart and liver primordia. In mid- and late-gestational stages, widely expressed.
Domain	Contains a cytoplasmic domain responsible for its TGN localization and recycling from the cell surface.
Post-translational modification	The inhibition peptide, which plays the role of an intramolecular chaperone, is autocatalytically removed in the endoplasmic reticulum (ER) and remains non-covalently bound to furin as a potent autoinhibitor. Following transport to the trans Golgi, a second cleavage within the inhibition propeptide results in propeptide dissociation and furin activation <span>By similarity</span> .  Phosphorylation is required for TGN localization of the endoprotease. In vivo, exists as di-, mono- and non-phosphorylated forms <span>By similarity</span> .
Sequence similarities	Belongs to the peptidase S8 family. Furin subfamily. Contains 1 homo B/P domain.

## Ontologies










### Keywords

Cellular component	Cell membrane Golgi apparatus Membrane
Domain	Signal Transmembrane Transmembrane helix
Ligand	Calcium Metal-binding
Molecular function	Hydrolase Protease Serine protease
PTM	Autocatalytic cleavage Cleavage on pair of basic residues Disulfide bond Glycoprotein Phosphoprotein Zymogen
Gene Ontology (GO)	
Biological process	aging Inferred from expression pattern. Source: RGD positive regulation of cell migration

Cellular component	Inferred from mutant phenotype. Source: RGD positive regulation of transforming growth factor beta receptor signaling pathway Inferred from mutant phenotype. Source: RGD
	Golgi cisterna Inferred from direct assay. Source: RGD endoplasmic reticulum membrane Inferred from direct assay. Source: RGD integral to membrane Inferred from electronic annotation. Source: UniProtKB-KW plasma membrane Inferred from direct assay. Source: RGD
Molecular function	metal ion binding Inferred from electronic annotation. Source: UniProtKB-KW serine-type endopeptidase activity Inferred from mutant phenotype. Source: RGD

Complete GO annotation...

Sequence annotation (Features)


Feature key	Position(s)	Length	Description	Graphical view	Feature id
Molecule processing					
	Signal peptide	1 – 24	24	Potential	
	Propeptide	25 – 107	83	Inhibition peptide By similarity	PRO_00000
	Chain	108 – 793	686	Furin	PRO_00000
Regions					
	Transmembrane	715 – 735	21	Helical; Potential	
	Region	758 – 761	4	Cell surface signal	
	Motif	498 – 500	3	Cell attachment site Potential	
	Motif	772 – 778	7	Trans Golgi network signal	
	Compositional bias	556 – 705	150	Cys-rich	
Sites					
	Active site	153	1	Charge relay system By similarity	

<input type="checkbox"/>	Active site	194	1	Charge relay system By similarity	I	
<input type="checkbox"/>	Active site	368	1	Charge relay system By similarity	I	
<input type="checkbox"/>	Metal binding	115	1	Calcium 1 By similarity	I	
<input type="checkbox"/>	Metal binding	162	1	Calcium 1 By similarity	I	
<input type="checkbox"/>	Metal binding	208	1	Calcium 1 By similarity	I	
<input type="checkbox"/>	Metal binding	258	1	Calcium 2 By similarity	I	
<input type="checkbox"/>	Metal binding	301	1	Calcium 2 By similarity	I	
<input type="checkbox"/>	Metal binding	331	1	Calcium 2 By similarity	I	
<input type="checkbox"/>	Site	75 – 76	2	Cleavage, second; by autolysis By similarity	I	
<input type="checkbox"/>	Site	107 – 108	2	Cleavage, first; by autolysis By similarity	I	

**Amino acid modifications**

<input type="checkbox"/>	Modified residue	772	1	Phosphoserine; by CK2 By similarity	I	
<input type="checkbox"/>	Modified residue	774	1	Phosphoserine; by CK2 By similarity	I	
<input type="checkbox"/>	Glycosylation	387	1	N-linked (GlcNAc...) Potential	I	
<input type="checkbox"/>	Glycosylation	440	1	N-linked (GlcNAc...) Potential	I	
<input type="checkbox"/>	Glycosylation	553	1	N-linked (GlcNAc...) Potential	I	
<input type="checkbox"/>	Disulfide bond	211 ↔ 360		By similarity	I I	
<input type="checkbox"/>	Disulfide bond	303 ↔ 333		By similarity	II	
<input type="checkbox"/>	Disulfide bond	450 ↔ 474		By similarity	II	

## Sequences

Sequence	Length	Mass (Da)
 P23377 [UniParc]. FASTA 793 86,653		
Last modified November 1, 1991. Version 1. Checksum: 87C22C345AE0A25C		
<div><div>102030405060</div><div>MELRPWLLWV VAAAGALVLL AAEARGQKIF TNTWAVHISG GPAVADSVAR KHGFHNLGQI</div></div> <div><div>708090100110120</div><div>FGDYHFWHR AVTKRSLSPH RPRHSRLQRV PQVKWLEQQV AKQRAKRDVY QEPTDPKFPQ</div></div> <div><div>130140150160170180</div><div>QWYLSGVTQR DLNVKEAWAQ GFTGRGIVVS ILDDGIEKNH PDLAGNYDPG ASFDVNDQDP</div></div> <div><div>190200210220230240</div><div>DPQPRYTQMN DNRHGTRCAG EVAAVANNGV CGVGVAYNAR IGGVRMLDGE VTDAVEARSL</div></div> <div><div>250260270280290300</div><div>GLNPNHIHIY SASWGPEDDG KTVDGPARLA EEAFRGVSQ GRGGLGSIFV WASGNGGREH</div></div> <div><div>310320330340350360</div><div>DSCNCDGYTN SIYTLSSISA TQFGNVPWYS EACSSTLATT YSSGNQNEKQ IVTTDLRQKC</div></div> <div><div>370380390400410420</div><div>TESHTGTSAS APLAAGIIAL TLEANKNLTW RDMQHLVVQT SKPAHLNAND WATNGVGRKV</div></div> <div><div>430440450460470480</div><div>SHSYGYGLLD AGAMVALAQN WTTVAPQRKC IIEILAEPKD IGKRLEVRKT VTACLGEPNH</div></div> <div><div>490500510520530540</div><div>ISRLEHVQAR LTLSYNRRGD LAIHLISPMG TRSTLLAARP HDYSADGFND WAFMTTHSWD</div></div> <div><div>550560570580590600</div><div>EDPSGEWVLE IENTSEANNY GTLTKFTLV L YGTASEGLSA PPESGCKTL TSSQACVVCE</div></div> <div><div>610620630640650660</div><div>EGFSLHQKSC VQRCPGFTP QVLDTHYTE NDVEIIRASV CTPCHASCAT CQGPAPTDCL</div></div> <div><div>670680690700710720</div><div>SCPSHASLDP VEQTCSRQSQ SSRESRPQQP PPALRPEVEV EPRLRAGLAS HLPEVLAGLS</div></div> <div><div>730740750760770780</div><div>CLIIALIFGI VFLFLHRCSG FSFRGVKVT MDRGLISYKG LPPEAWQEEC PSDSEDEGR</div></div> <div><div>790</div><div>GERTAFIKDQ SAL</div></div>		

« Hide

## References

- [1] "Sequence of the cDNA encoding rat furin, a possible propeptide-processing endoprotease."  
Misumi Y., Sohoda M., Ikehara Y.  
Nucleic Acids Res. 18:6719-6719(1990) [PubMed: 2251148] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA].  
Strain: Wistar.  
Tissue: Liver.
- [2] "PACS-1 defines a novel gene family of cytosolic sorting proteins required for trans-Golgi network localization."  
Wan L., Molloy S.S., Thomas L., Liu G., Xiang Y., Rybak S.L., Thomas G.  
Cell 94:205-216(1998) [PubMed: 9695949] [Abstract]  
Cited for: INTERACTION WITH PACS1.  
Tissue: Brain.



+ Additional computationally mapped references.

## Cross-references

### Sequence databases

<input checked="" type="radio"/> EMBL <input checked="" type="radio"/> GenBank <input checked="" type="radio"/> DDBJ	X55660 mRNA. Translation: CAA39193.1.
IPI	IPI00210230.
PIR	KXRTF. S13106.
RefSeq	NP_062204.1. NM_019331.1.
UniGene	Rn.3220.

### 3D structure databases

ProteinModelPortal	P23377.
SMR	P23377. Positions 30-99, 108-574, 592-665.
ModBase	Search...

### Protein-protein interaction databases

STRING	P23377.
--------	---------

### Protein family/group databases

MEROPS	S08.071.
--------	----------

### Proteomic databases

PRIDE	P23377.
-------	---------

### Genome annotation databases

Ensembl	ENSRNOT00000015521; ENSRNOP00000015521; ENSRNOG00000011352.
GeneID	54281.
KEGG	rno:54281.
UCSC	NM_019331. rat.

### Organism-specific databases

CTD	54281.
RGD	3274. Furin.

### Phylogenomic databases

eggNOG	maNOG16185.
GeneTree	ENSGT00600000084064.
HOVERGEN	HBG008705.

InParanoid	P23377.
OrthoDB	EOG4ZW59M.
PhylomeDB	P23377.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.21.75. 248.
<b>Gene expression databases</b>	
ArrayExpress	P23377.
Genevestigator	P23377.
GermOnline	ENSRNOG00000011352. Rattus norvegicus.
<b>Family and domain databases</b>	
InterPro	IPR006212. Furin_repeat. IPR008979. Galactose-bd-like. IPR009030. Growth_fac_rcpt. IPR000209. Peptidase_S8/S53. IPR022398. Peptidase_S8/S53_AS. IPR015500. Peptidase_S8_subtilisin-rel. IPR009020. Prot_inh_propept. IPR002884. PrprotnconvertsP. [Graphical view]
Gene3D	G3DSA:3.40.50.200. Pept_S8_S53. 1 hit.
PANTHER	PTHR10795. SubtilSerProt. 1 hit.
Pfam	PF01483. P_proprotein. 1 hit. PF00082. Peptidase_S8. 1 hit. [Graphical view]
PRINTS	PR00723. SUBTILISIN.
SMART	SM00261. FU. 2 hits. [Graphical view]
SUPFAM	SSF49785. Gal_bind_like. 1 hit. SSF57184. Grow_fac_recept. 1 hit. SSF52743. Pept_S8_S53. 1 hit. SSF54897. Prot_inh_propept. 1 hit.
PROSITE	PS00136. SUBTILASE_ASP. 1 hit. PS00137. SUBTILASE_HIS. 1 hit. PS00138. SUBTILASE_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
NextBio	610854.
<b>Entry information</b>	
Entry name	FURIN_RAT
Accession	Primary (citable) accession number: <b>P23377</b>

Entry history	Integrated into November 1, 1991 UniProtKB/Swiss-Prot: Last sequence November 1, 1991 update: Last modified: March 8, 2011 This is version 110 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Relevant documents</b>	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact



**P29119 (FURI1\_XENLA) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified October 5, 2010. Version 81.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Furin-1</b> EC=3.4.21.75 <i>Alternative name(s):</i> Dibasic-processing enzyme Paired basic amino acid residue-cleaving enzyme Short name=PACE
Gene names	Name: <b>furin</b>
Organism	<b><i>Xenopus laevis</i> (African clawed frog)</b>
Taxonomic identifier	8355 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Amphibia › Batrachia › Anura › Mesobatrachia › Pipoidae › Pipidae › Xenopodinae › Xenopus › Xenopus

**Protein attributes**

Sequence length	783 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at transcript level.

**General annotation (Comments)**

Function	Furin is likely to represent the ubiquitous endoprotease activity within constitutive secretory pathways and capable of cleavage at the RX(K/R)R consensus motif.
Catalytic activity	Release of mature proteins from their proproteins by cleavage of -Arg-Xaa-Yaa-Arg- -Zaa- bonds, where Xaa can be any amino acid and Yaa is Arg or Lys. Releases albumin, complement component C3 and vWF from their respective precursors.
Subcellular location	Membrane; Single-pass membrane protein <span>Potential</span> .
Tissue specificity	In all tissues analyzed.
Sequence similarities	Belongs to the peptidase S8 family. Furin subfamily.









**Ontologies**

<b>Keywords</b>	
Cellular component	Membrane



Domain	Signal Transmembrane Transmembrane helix
Molecular function	Hydrolase Protease Serine protease
PTM	Cleavage on pair of basic residues Disulfide bond Glycoprotein Zymogen
<b>Gene Ontology (GO)</b>	
Biological process	proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	integral to membrane Inferred from electronic annotation. Source: UniProtKB-KW
Molecular function	serine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro
Complete GO annotation...	

**Sequence annotation (Features)**

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
<b>Molecule processing</b>					
 Signal peptide	1 – 24	24	Potential		
 Propeptide	25 – 105	81	Potential		PRO_0000027034
 Chain	106 – 783	678	Furin-1		PRO_0000027035
<b>Regions</b>					
 Transmembrane	702 – 725	24	Helical; Potential		
 Region	118 – 418	301	Catalytic		
<b>Sites</b>					
 Active site	151	1	Charge relay system By similarity		
 Active site	192	1	Charge relay system By similarity		
 Active site	366	1	Charge relay		

## Sequences

<http://www.uniprot.org/uniprot/P29119>

## References

- [1] "Prohormone processing in *Xenopus* oocytes: characterization of cleavage signals and cleavage enzymes."  
 Korner J., Chun J., O'Bryan L., Axel R.  
 Proc. Natl. Acad. Sci. U.S.A. 88:11393-11397(1991) [PubMed: 1722329] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA].

## Cross-references

### Sequence databases

<input checked="" type="radio"/> EMBL <input checked="" type="radio"/> GenBank <input checked="" type="radio"/> DDBJ	M80471 mRNA. Translation: AAA49717.1.
--	---------------------------------------

PIR	A41627. B41627.
-----	--------------------

UniGene	XI.788.
---------	---------

### 3D structure databases

ProteinModelPortal	P29119.
--------------------	---------

SMR	P29119. Positions 28-97, 108-572, 590-665.
-----	--

ModBase	Search...
---------	-----------

### Protein family/group databases

MEROPS	S08.071.
--------	----------

### Phylogenomic databases

HOVERGEN	HBG008705.
----------	------------

### Enzyme and pathway databases

BRENDA	3.4.21.75. 648.
--------	-----------------

### Family and domain databases

InterPro	IPR006212. Furin_repeat. IPR008979. Galactose-bd-like. IPR009030. Growth_fac_rcpt. IPR000209. Peptidase_S8/S53. IPR022398. Peptidase_S8/S53_AS. IPR015500. Peptidase_S8_subtilisin-rel. IPR009020. Prot_inh_propept. IPR002884. PrprotncnvertsP. [Graphical view]
----------	---

Gene3D	G3DSA:3.40.50.200. Pept_S8_S53. 1 hit.
--------	--

PANTHER	PTHR10795. SubtilSerProt. 1 hit.
---------	----------------------------------

Pfam	PF01483. P_proprotein. 1 hit. PF00082. Peptidase_S8. 1 hit. [Graphical view]
------	--

PRINTS	PR00723. SUBTILISIN.
SMART	SM00261. FU. 2 hits. [Graphical view]
SUPFAM	SSF49785. Gal_bind_like. 1 hit. SSF57184. Grow_fac_recept. 1 hit. SSF52743. Pept_S8_S53. 1 hit. SSF54897. Prot_inh_propept. 1 hit.
PROSITE	PS00136. SUBTILASE_ASP. 1 hit. PS00137. SUBTILASE_HIS. 1 hit. PS00138. SUBTILASE_SER. 1 hit. [Graphical view]
ProtoNet	Search...

### Entry information

Entry name	FURI1_XENLA
Accession	Primary (citable) accession number: <b>P29119</b>
Entry history	Integrated into      December 1, 1992 UniProtKB/Swiss- Prot: Last sequence      December 1, 1992 update: Last modified:      October 5, 2010 This is version 81 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program

### Relevant documents

Peptidase families  
Classification of peptidase families and list of entries

SIMILARITY comments  
Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact





**P25779 (CYSP\_TRYCR) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified January 11, 2011. Version 91.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Cruzipain</b> EC=3.4.22.51 <i>Alternative name(s):</i> Cruzaine Major cysteine proteinase
Organism	<b>Trypanosoma cruzi</b>
Taxonomic identifier	5693 [NCBI]
Taxonomic lineage	Eukaryota › Euglenozoa › Kinetoplastida › Trypanosomatidae › Trypanosoma › Schizotrypanum

**Protein attributes**










Sequence length	467 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Hydrolyzes chromogenic peptides at the carboxyl Arg or Lys; requires at least one more amino acid, preferably Arg, Phe, Val or Leu, between the terminal Arg or Lys and the amino-blocking group.  The cysteine protease may play an important role in the development and differentiation of the parasites at several stages of their life cycle.
Catalytic activity	Broad endopeptidase specificity similar to that of cathepsin L.
Enzyme regulation	Strongly inhibited by E-64 (L-trans-epoxysuccinylleucylamido(4-guanidino)butane), Leupeptin, and N-alpha-p-tosyl-L-lysine chloromethyl ketone.
Developmental stage	Present in all developmental stages.
Miscellaneous	Purified cruzipain is able to degrade itself, yielding a complex mixture of small peptides, and a major 25 kDa fragment.
Sequence similarities	Belongs to the peptidase C1 family.

**Ontologies****Keywords**

Domain	Signal
--------	--------

Molecular function	Hydrolase Protease Thiol protease					
PTM	Autocatalytic cleavage Disulfide bond Glycoprotein Zymogen					
Technical term	3D-structure Direct protein sequencing					
Gene Ontology (GO)						
Biological process	proteolysis Inferred from electronic annotation. Source: InterPro					
Molecular function	cysteine-type endopeptidase activity Inferred from electronic annotation. Source: InterPro					
Complete GO annotation...						
Sequence annotation (Features)						
Feature key	Position(s)	Length	Description	Graphical view	Feature identifier	
Molecule processing						
	Signal peptide	1 – 18	18	<div>Probable</div>		
	Propeptide	19 – 122	104	Activation peptide <div>Probable</div>		PRO_000002637
	Chain	123 – 467	345	Cruzipain		PRO_000002637
Sites						
	Active site	147	1			
	Active site	284	1			
	Active site	304	1			
	Site	337 – 338	2	Cleavage; by autolysis <div>Ref.1</div>		
Amino acid modifications						
	Glycosylation	169	1	N-linked (GlcNAc...) <div>Potential</div>		
	Glycosylation	292	1	N-linked (GlcNAc...) <div>Potential</div>		

<input type="checkbox"/>	Glycosylation	377	1	N-linked (GlcNAc...) Potential		
<input type="checkbox"/>	Disulfide bond	144 ↔ 185				
<input type="checkbox"/>	Disulfide bond	178 ↔ 223				
<input type="checkbox"/>	Disulfide bond	277 ↔ 325				
<b>Natural variations</b>						
<input type="checkbox"/>	Natural variant	9	1	L → S in clone 1800 -2.		
<input type="checkbox"/>	Natural variant	35	1	T → A in clone 1800 -2.		
<input type="checkbox"/>	Natural variant	39	1	A → V in clone 1800 -2.		
<input type="checkbox"/>	Natural variant	51	1	S → N in clone 1800 -4.		
<input type="checkbox"/>	Natural variant	56	1	A → R in clone 1800 -2.		
<input type="checkbox"/>	Natural variant	76	1	N → G in clone 1800 -4.		
<input type="checkbox"/>	Natural variant	109	1	Q → G in clone 1800 -4.		
<input type="checkbox"/>	Natural variant	117	1	K → N in clone 1800 -2.		
<input type="checkbox"/>	Natural variant	146	1	S → G		
<input type="checkbox"/>	Natural variant	157 – 158	2	EC → SG in strain: RA.		
<input type="checkbox"/>	Natural variant	186	1	S → G in clones 1800-2 and 1800- 4.		
<input type="checkbox"/>	Natural variant	204	1	A → G in strain: RA.		
<input type="checkbox"/>	Natural variant	250 – 251	2	WL → CV in clones 1800-2 and 1800- 4.		

<input type="checkbox"/>	Natural variant	261 – 262	2	VD → H in strain: RA.		
<input type="checkbox"/>	Natural variant	286	1	V → F in clone 1800 -4.		
<input type="checkbox"/>	Natural variant	286	1	V → L in strain: RA.		
<input type="checkbox"/>	Natural variant	308	1	T → A in clone 1800 -2.		
<input type="checkbox"/>	Natural variant	313	1	E → D		
<input type="checkbox"/>	Natural variant	409	1	L → F in clone 1800 -2.		
<input type="checkbox"/>	Natural variant	427	1	K → Q in clone 1800 -2.		
<input type="checkbox"/>	Natural variant	430	1	R → W in clone 1800 -2.		
<input type="checkbox"/>	Natural variant	457	1	H → Y in clone 1800 -2.		
<input type="checkbox"/>	Natural variant	461	1	H → Q in clone 1800 -2.		

Experimental info

<input type="checkbox"/>	Sequence conflict	357	1	S → C Ref.4		
--------------------------	-------------------	-----	---	----------------	--	--

Secondary structure

1. ....

46

■ Helix ■ Strand ■ Turn

Details...

Sequences

Sequence	Length	Mass (Da)
<input type="checkbox"/> P25779 [UniParc]. Last modified May 1, 1992. Version 1. Checksum: B93EBA49B511363D	FASTA 467	49,836
<div><div><div>102030405060</div><div>MSGWARALLLA AVLVMACLVPAATASLHA EETLTSQFAE FKQKHGRVYE SAAEEAFRLS</div></div><div><div>708090100110120</div><div>VFRENLFLLAR LHAAANPHAT FGVTPFSDLT REEFRSRYHN GAAHFAAAQE RARVPVKVEV</div></div></div>		



```

      130      140      150      160      170      180
VGAPAAVDWR ARGAVTAVKD QGQCGSCWAF SAIGNVECQW FLAGHPLTNL SEQMLVSCDK

      190      200      210      220      230      240
TDSGCSGGLM NNAFEWIVQE NNGAVYTEDS YPYASGEGIS PPCTTSGHTV GATITGHVEL

      250      260      270      280      290      300
PQDEAQIAAW LAVNGPVAVA VDASSWMTYT GGVMTSCVSE QLDHGVLLVG YNDSAAVPYW

      310      320      330      340      350      360
IIKNSWTTQW GEEGYIRIAK GSNQCLVKEE ASSAVVGGPG PTPEPTTTT TSAPGPSPSY

      370      380      390      400      410      420
FVQMSCTDAA CIVGCENVTL PTGQCLLTTS GVSAIVTCGA ETLTEEVFLT STHCSGPSVR

      430      440      450      460
SSVPLNKCNR LLRGSVEFFC GSSSSGRLAD VDRQRRHQPY HSRHRRL

```

« Hide

## References

- [1] **"The sequence, organization, and expression of the major cysteine protease (cruzain) from Trypanosoma cruzi."**  
Eakin A.E., Mills A.A., Harth G., McKerrow J.H., Craik C.S.  
J. Biol. Chem. 267:7411-7420(1992) [PubMed: 1559982] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE.  
Strain: Tulahuen.
- [2] **"The major cysteine proteinase (cruzipain) from Trypanosoma cruzi is encoded by multiple polymorphic tandemly organized genes located on different chromosomes."**  
Campetella O., Henriksson J., Aaslund L., Frasch A.C.C., Pettersson U., Cazzulo J.J.  
Mol. Biochem. Parasitol. 50:225-234(1992) [PubMed: 1311053] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [GENOMIC DNA] (CLONES 1800-2 AND 1800-4).  
Strain: Tulahuen 2.
- [3] **"Amplification and sequencing of genomic DNA fragments encoding cysteine proteases from protozoan parasites."**  
Eakin A.E., Bouvier J., Sakanari J.A., Craik C.S., McKerrow J.H.  
Mol. Biochem. Parasitol. 39:1-8(1990) [PubMed: 2406590] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 141-306.  
Strain: RA.
- [4] **"The C-terminal extension of the major cysteine proteinase (cruzipain) from Trypanosoma cruzi."**  
Aaslund L., Henriksson J., Campetella O., Frasch A.C.C., Pettersson U., Cazzulo J.J.  
Mol. Biochem. Parasitol. 45:345-348(1991) [PubMed: 2038364] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA] OF 295-467.  
Strain: Tulahuen 2.
- [5] **"Self-proteolysis of the cysteine proteinase, cruzipain, from Trypanosoma cruzi gives a major fragment corresponding to its carboxy-terminal domain."**  
Hellman U., Wernstedt C., Cazzulo J.J.  
Mol. Biochem. Parasitol. 44:15-21(1991) [PubMed: 2011151] [Abstract]  
Cited for: AUTOCATALYSIS OF C-TERMINAL.  
Strain: Tulahuen 2.
- [6] **"Some kinetic properties of a cysteine proteinase (cruzipain) from Trypanosoma cruzi."**  
Cazzulo J.J., Cazzulo-Franke M.C., Martinez J., Franke de Cazzulo B.M.  
Biochim. Biophys. Acta 1037:186-191(1990) [PubMed: 2407295] [Abstract]  
Cited for: SPECIFICITY.  
Strain: Tulahuen 2.

- [7] **"Further characterization and partial amino acid sequence of a cysteine proteinase from Trypanosoma cruzi."**  
Cazzulo J.J., Couso R., Raimondi A., Wernstedt C., Hellman U.  
Mol. Biochem. Parasitol. 33:33-42(1989) [PubMed: 2651912] [Abstract]  
Cited for: PROTEIN SEQUENCE OF 123-146 AND 304-317.
- [8] **"Structural determinants of specificity in the cysteine protease cruzain."**  
Gillmor S.A., Craik C.S., Fletterick R.J.  
Protein Sci. 6:1603-1611(1997) [PubMed: 9260273] [Abstract]  
Cited for: X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 123-337.
- + Additional computationally mapped references.

## Cross-references

### Sequence databases

<ul style="list-style-type: none"> <li>⊙ EMBL</li> <li>⊙ GenBank</li> <li>⊙ DDBJ</li> </ul>	M84342 Genomic DNA. Translation: AAA30181.1. M69121 Genomic DNA. Translation: AAA30269.1. M69121 Genomic DNA. Translation: AAA30270.1. X54414 mRNA. Translation: CAA38278.1. M27305 Genomic DNA. Translation: AAA30180.1.
---	---

PIR	A60667. S16162.
-----	--------------------

### 3D structure databases

<ul style="list-style-type: none"> <li>⊙ PDBe</li> <li>⊙ RCSB PDB</li> <li>⊙ PDBj</li> </ul>	Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
	1AIM	X-ray	2.00	A	123-337	[»]
	1EWL	X-ray	2.00	A	123-337	[»]
	1EWM	X-ray	2.00	A	123-337	[»]
	1EWO	X-ray	2.10	A	123-337	[»]
	1EWP	X-ray	1.75	A	123-337	[»]
	1F29	X-ray	2.15	A/B/C	123-337	[»]
	1F2A	X-ray	1.60	A	123-337	[»]
	1F2B	X-ray	1.80	A	123-337	[»]
	1F2C	X-ray	2.00	A	123-337	[»]
	1ME3	X-ray	1.20	A	123-337	[»]
	1ME4	X-ray	1.20	A	123-337	[»]
	1U9Q	X-ray	2.30	X	123-337	[»]
	2AIM	X-ray	2.20	A	123-337	[»]
	2OZ2	X-ray	1.95	A/C	123-337	[»]
	3HD3	X-ray	1.75	A/B	123-337	[»]
	3I06	X-ray	1.10	A	123-337	[»]
	3IUT	X-ray	1.20	A	123-337	[»]
	3KKU	X-ray	1.28	A	123-337	[»]
	3LXS	X-ray	1.50	A/C	123-337	[»]

ProteinModelPortal	P25779.
--------------------	---------

SMR	P25779. Positions 31-337.
-----	---------------------------

ModBase	Search...
---------	-----------

### Protein family/group databases

MEROPS	C01.075.
--------	----------

**Enzyme and pathway databases**

BRENDA	3.4.22.51. 884.
--------	-----------------

**Family and domain databases**

InterPro	IPR021981. DUF3586. IPR000169. Pept_cys_AS. IPR013128. Peptidase_C1A. IPR000668. Peptidase_C1A_C. IPR013201. Prot_inhib_I29. [Graphical view]
----------	--

PANTHER	PTHR12411. Peptidase_C1A. 1 hit.
---------	----------------------------------

Pfam	PF12131. DUF3586. 1 hit. PF08246. Inhibitor_I29. 1 hit. PF00112. Peptidase_C1. 1 hit. [Graphical view]
------	---

PRINTS	PR00705. PAPAIN.
--------	------------------

SMART	SM00848. Inhibitor_I29. 1 hit. SM00645. Pept_C1. 1 hit. [Graphical view]
-------	--

PROSITE	PS00640. THIOL_PROTEASE_ASN. 1 hit. PS00139. THIOL_PROTEASE_CYS. 1 hit. PS00639. THIOL_PROTEASE_HIS. 1 hit. [Graphical view]
---------	---

ProtoNet	Search...
----------	-----------

**Entry information**

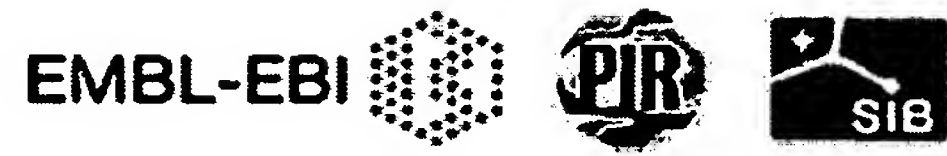
Entry name	CYSP_TRYCR
Accession	Primary (citable) accession number: <b>P25779</b>
Entry history	Integrated into May 1, 1992 UniProtKB/Swiss-Prot: Last sequence May 1, 1992 update: Last modified: January 11, 2011 This is version 91 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)

**Relevant documents**

PDB cross-references Index of Protein Data Bank (PDB) cross-references
---

Peptidase families Classification of peptidase families and list of entries
--

SIMILARITY comments Index of protein domains and families
--





## Protein

Translations of Life

Display Settings: GenPept

## cysteine proteinase [Trypanosoma cruzi]

GenBank: AAA30269.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAA30269 249 aa linear INV 26-APR-1993  
DEFINITION cysteine proteinase [Trypanosoma cruzi].  
ACCESSION AAA30269  
VERSION AAA30269.1 GI:162334  
DBSOURCE locus TRBTUL2TAN accession M69121.1  
KEYWORDS  
SOURCE Trypanosoma cruzi  
ORGANISM Trypanosoma cruzi  
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;  
Trypanosoma; Schizotrypanum.  
REFERENCE 1 (residues 1 to 249)  
AUTHORS Campetella, O., Henriksson, J., Aslund, L., Frasch, A.C.C.,  
Pettersson, J. and Cazzulo, J.J.  
TITLE The major cysteine proteinase (cruzipain) from Trypanosoma cruzi is  
encoded by genes organized in tandems located in different  
chromosomes  
JOURNAL Unpublished  
COMMENT Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..249  
/organism="Trypanosoma cruzi"  
/strain="TUL 2"  
/db\_xref="taxon:5693"  
Protein 1..249  
/product="cysteine proteinase"  
mat peptide <1..249  
/product="cysteine proteinase"  
Region <5..116  
/region\_name="Peptidase C1A"  
/note="Peptidase C1A subfamily (MEROPS database  
nomenclature); composed of cysteine peptidases (CPs)  
similar to papain, including the mammalian CPs (cathepsins  
B, C, F, H, L, K, O, S, V, X and W). Papain is an  
endopeptidase with specific substrate preferences...;  
cd02248"  
/db\_xref="CDD:30292"  
Site order(42,64,67,112)  
/site\_type="other"  
/note="S2 subsite"  
/db\_xref="CDD:30292"  
Site order(66,86)  
/site\_type="active"  
/db\_xref="CDD:30292"  
Region <142..199  
/region\_name="DUF3586"  
/note="Protein of unknown function (DUF3586); pfam12131"  
/db\_xref="CDD:152566"  
CDS 1..249  
/coded\_by="M69121.1:<1..750"  
ORIGIN  
1 isppcttsgh tvgatitghv elpqdeaqia acvavngpva vavdasswmt ytggvmtscv  
61 seqldhgvll vgyndsaaav ywiiknswta qwgedgyiri akgsnqclvk eeassavvgg  
121 pgptpepttt tttsapgpsp syfvqmsctd aacivgcenv tlptggcllt tsgvsaivtc  
181 gaetlteevf ftstchsgps vrssvplnqc nwlirgsvef fcgssssgrl advdrqrryq  
241 pyqsrhrrl  
//

## Protein

Translations of Life

Display Settings: GenPept

## cysteine proteinase [Trypanosoma cruzi]

GenBank: CAA38278.1

[FASTA](#) [Graphics](#)Go to:

LOCUS CAA38278 173 aa linear INV 19-APR-1991  
DEFINITION cysteine proteinase [Trypanosoma cruzi].  
ACCESSION CAA38278  
VERSION CAA38278.1 GI:10606  
DBSOURCE embl accession X54414.1  
KEYWORDS  
SOURCE Trypanosoma cruzi  
ORGANISM Trypanosoma cruzi  
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;  
Trypanosoma; Schizotrypanum.  
REFERENCE 1  
AUTHORS Aslund,L., Henriksson,J., Campetella,O., Frasch,A.C., Pettersson,U.  
and Cazzulo,J.J.  
TITLE The C-terminal extension of the major cysteine proteinase  
(cruzipain) from Trypanosoma cruzi  
JOURNAL Mol. Biochem. Parasitol. 45 (2), 345-347 (1991)  
PUBMED 2038364  
REFERENCE 2 (residues 1 to 173)  
AUTHORS Aslund,L.  
TITLE Direct Submission  
JOURNAL Submitted (13-AUG-1990) Dept. Medical Genetics, Biomedical Center,  
Box 589, S-75123 Uppsala, Sweden  
FEATURES  
Location/Qualifiers  
source 1..173  
/organism="Trypanosoma cruzi"  
/db\_xref="taxon:5693"  
Protein <1..173  
/product="cysteine proteinase"  
Region <3..40  
/region\_name="Peptidase\_C1"  
/note="C1 Peptidase family (MEROPS database nomenclature),  
also referred to as the papain family; composed of two  
subfamilies of cysteine peptidases (CPs), C1A (papain) and  
C1B (bleomycin hydrolase). Papain-like enzymes are mostly  
endopeptidases with some...; cl00298"  
/db\_xref="CDD:153664"  
Region <66..123  
/region\_name="DUF3586"  
/note="Protein of unknown function (DUF3586); pfam12131"  
/db\_xref="CDD:152566"  
CDS 1..173  
/coded\_by="X54414.1:<1..522"  
/db\_xref="GOA:P25779"  
/db\_xref="InterPro:IPR000169"  
/db\_xref="InterPro:IPR000668"  
/db\_xref="InterPro:IPR013128"  
/db\_xref="InterPro:IPR013201"  
/db\_xref="InterPro:IPR021981"  
/db\_xref="PDB:1AIM"  
/db\_xref="PDB:1EWL"  
/db\_xref="PDB:1EWM"  
/db\_xref="PDB:1EWC"  
/db\_xref="PDB:1EWP"  
/db\_xref="PDB:1F29"  
/db\_xref="PDB:1F2A"  
/db\_xref="PDB:1F2B"  
/db\_xref="PDB:1F2C"  
/db\_xref="PDB:1ME3"  
/db\_xref="PDB:1ME4"  
/db\_xref="PDB:1U9Q"  
/db\_xref="PDB:2AIM"  
/db\_xref="PDB:2EFM"  
/db\_xref="PDB:2O22"  
/db\_xref="PDB:3HD3"  
/db\_xref="PDB:3I06"  
/db\_xref="PDB:3IUT"  
/db\_xref="PDB:3KKU"  
/db\_xref="UniProtKB/Swiss-Prot:P25779"  
ORIGIN  
1 aavpywiikn swtaqwgcdg yiriakgsnq clvkeeassa vvggpgptpe ptttttsap  
61 gspspysfvqm sctdaacivg cenvtlptgq clttsgvsa ivtcgaetlt eevfftsthc  
121 sgpsvrssvp lncnrlrrg sveffcgsss sgrladvdrq rryqpyhsrh rrl  
//

## Protein

Translations of Life

Display Settings: GenPept

## cysteine proteinase [Trypanosoma cruzi]

GenBank: AAA30270.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAA30270 218 aa linear INV 26-APR-1993.  
DEFINITION cysteine proteinase [Trypanosoma cruzi].  
ACCESSION AAA30270  
VERSION AAA30270.1 GI:162335  
DBSOURCE locus TRBTUL2TAN accession M69121.1  
KEYWORDS  
SOURCE Trypanosoma cruzi  
ORGANISM Trypanosoma cruzi  
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;  
Trypanosoma; Schizotrypanum.  
REFERENCE 1 (residues 1 to 218)  
AUTHORS Campetella, O., Henriksson, J., Aslund, L., Frasch, A.C.C.,  
Pettersson, J. and Cazzulo, J.J.  
TITLE The major cysteine proteinase (cruzipain) from Trypanosoma cruzi is  
encoded by genes organized in tandems located in different  
chromosomes  
JOURNAL Unpublished  
COMMENT Method: conceptual translation.  
FEATURES  
Location/Qualifiers  
source 1..218  
/organism="Trypanosoma cruzi"  
/strain="TUL 2"  
/db\_xref="taxon:5693"  
Protein 1..218  
/product="cysteine proteinase"  
sig\_peptide 1..18  
/note="pre-enzyme domain"  
Region 38..93  
/region\_name="Inhibitor\_I29"  
/note="Cathepsin propeptide inhibitor domain (I29);  
cl07031"  
/db\_xref="CDD:157580"  
Region 124..>218  
/region\_name="Peptidase\_C1A"  
/note="Peptidase C1A subfamily (MEROPS database  
nomenclature); composed of cysteine peptidases (CPs)  
similar to papain, including the mammalian CPs (cathepsins  
B, C, F, H, L, K, O, S, V, X and W). Papain is an  
endopeptidase with specific substrate preferences...;  
cd02248"  
/db\_xref="CDD:30292"  
mat\_peptide 125..>218  
/product="cysteine proteinase"  
/note="pre-enzyme domain"  
Site order(141,147)  
/site\_type="active"  
/db\_xref="CDD:30292"  
CDS 1..218  
/coded\_by="M69121.1:1192..>1845"  
ORIGIN  
1 msgwaralsl aavlvvmacl vpaataslha eetlasqfve fkqkhgrvye saaeerfrls  
61 vfrenlflar lhaaanphat fgvtpfsdlt reefrsryhn gaahfaaaqe rarvpvnvev  
121 vgapaavdwr argavtavkd qgqcgscwaf saignvecqw flaghpltnl seqmlvscdk  
181 tdsgcggglm nnafewivqe nngavyteds ypyasgeg  
//

## Protein

Translations of Life

Display Settings: GenPept

## cysteine protease [Trypanosoma cruzi]

GenBank: AAA30180.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAA30180 165 aa linear INV 26-APR-1993  
DEFINITION cysteine protease [Trypanosoma cruzi].  
ACCESSION AAA30180  
VERSION AAA30180.1 GI:162044  
DBSOURCE locus TRBCYPAA accession M27305.1  
KEYWORDS  
SOURCE Trypanosoma cruzi  
ORGANISM Trypanosoma cruzi  
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;  
Trypanosoma; Schizotrypanum.  
REFERENCE 1 (residues 1 to 165)  
AUTHORS Eakin,A.E.  
JOURNAL Unpublished  
REFERENCE 2 (sites)  
AUTHORS Eakin,A.E., Bouvier,J., Sakanari,J.A., Craik,C.S. and McKerrow,J.H.  
TITLE Amplification and sequencing of genomic DNA fragments encoding  
cysteine proteases from protozoan parasites  
JOURNAL Mol. Biochem. Parasitol. 39 (1), 1-8 (1990)  
PUBMED 2406590  
COMMENT [2] sites; for [1].  
Draft entry and computer-readable sequence for [1] kindly submitted  
by A.E.Eakin, 24-AUG-1989.  
Method: conceptual translation.  
FEATURES  
source Location/Qualifiers  
1..165  
/organism="Trypanosoma cruzi"  
/db\_xref="taxon:5693"  
Protein 1..165  
/name="cysteine protease"  
Region 1..165  
/region\_name="Peptidase\_C1A"  
/note="Peptidase C1A subfamily (MEROPS database  
nomenclature); composed of cysteine peptidases (CPs)  
similar to papain, including the mammalian CPs (cathepsins  
B, C, F, H, L, K, O, S, V, X and W). Papain is an  
endopeptidase with specific substrate preferences...;  
cd02248"  
/db\_xref="CDD:30292"  
Site order(1,7,143,163)  
/site\_type="active"  
/db\_xref="CDD:30292"  
Site order(49..50,120,141,144)  
/site\_type="other"  
/note="S2 subsite"  
/db\_xref="CDD:30292"  
CDS 1..165  
/coded\_by="M27305.1:<1..>495"  
ORIGIN  
1 qgqcgscwaf saignvsgqw flaghpltnl seqmlvscdk tdsgcsgglm nnafewivqe  
61 nnggvtyeds ypyasgegis ppcttsghtv gatitghvel pqdeaqlaaw lavngpvava  
121 hasswmtytg gvmtscvseq ldhglllvgy ndsaavpywi iknsw  
//



**P00749 (UROK\_HUMAN) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 158.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Urokinase-type plasminogen activator</b> Short name=U-plasminogen activator Short name=uPA EC=3.4.21.73 <i>Cleaved into the following 3 chains:</i> 1. <b><i>Urokinase-type plasminogen activator long chain A</i></b> 2. <b><i>Urokinase-type plasminogen activator short chain A</i></b> 3. <b><i>Urokinase-type plasminogen activator chain B</i></b>
Gene names	Name: <b>PLAU</b>
Organism	<b>Homo sapiens (Human)</b> [Complete proteome]
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo

**Protein attributes**

Sequence length	431 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Specifically cleave the zymogen plasminogen to form the active enzyme plasmin.
Catalytic activity	Specific cleavage of Arg- -Val bond in plasminogen to form plasmin.
Subunit structure	Found in high and low molecular mass forms. Each consists of two chains, A and B. The high molecular mass form contains a long chain A which is cleaved to yield a short chain A. Binds LRP1B; binding is followed by internalization and degradation. Interacts with MRC2. Interacts with PLAUR. <a href="#">Ref.18</a> <a href="#">Ref.19</a>
Subcellular location	Secreted.
Tissue specificity	Expressed in the prostate gland and prostate cancers. <a href="#">Ref.20</a>
Post-translational modification	Phosphorylation of Ser-158 and Ser-323 abolishes proadhesive ability but does not interfere with receptor binding.


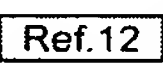
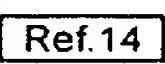




Pharmaceutical use	Available under the name Abbokinase (Abbott). Used in Pulmonary Embolism (PE) to initiates fibrinolysis. Clinically used for therapy of thrombolytic disorders.
Sequence similarities	<p>Belongs to the peptidase S1 family.</p> <p>Contains 1 EGF-like domain.</p> <p>Contains 1 kringle domain.</p> <p>Contains 1 peptidase S1 domain.</p>
<b>Ontologies</b>	
<b>Keywords</b>	
Biological process	<p>Blood coagulation</p> <p>Fibrinolysis</p> <p>Plasminogen activation</p>
Cellular component	Secreted
Coding sequence diversity	<p>Alternative splicing</p> <p>Polymorphism</p>
Domain	<p>EGF-like domain</p> <p>Kringle</p> <p>Signal</p>
Molecular function	<p>Hydrolase</p> <p>Protease</p> <p>Serine protease</p>
PTM	<p>Disulfide bond</p> <p>Glycoprotein</p> <p>Phosphoprotein</p> <p>Zymogen</p>
Technical term	<p>3D-structure</p> <p>Complete proteome</p> <p>Direct protein sequencing</p> <p>Pharmaceutical</p>
<b>Gene Ontology (GO)</b>	
Biological process	<p>blood coagulation</p> <p>Traceable author statement. Source: Reactome</p> <p>chemotaxis</p> <p>Traceable author statement. Source: ProtInc</p> <p>fibrinolysis</p> <p>Traceable author statement. Source: Reactome</p> <p>proteolysis</p> <p>Traceable author statement. Source: ProtInc</p> <p>regulation of cell adhesion mediated by integrin</p> <p>Inferred from direct assay. Source: BHF-UCL</p> <p>regulation of receptor activity</p> <p>Inferred from direct assay. Source: BHF-UCL</p> <p>regulation of smooth muscle cell migration</p> <p>Inferred from direct assay. Source: BHF-UCL</p> <p>regulation of smooth muscle cell-matrix adhesion</p>

Cellular component	Inferred from direct assay. Source: BHF-UCL signal transduction Traceable author statement. Source: ProtInc
	cell surface Inferred from direct assay. Source: BHF-UCL extracellular space Inferred from direct assay. Source: BHF-UCL plasma membrane Inferred from experiment. Source: Reactome
Molecular function	serine-type endopeptidase activity Traceable author statement. Source: Reactome
Complete GO annotation...	






Alternative products

This entry describes 2 isoforms produced by <b>alternative splicing</b> . [Align] [Select]
<b>Isoform 1 (identifier: P00749-1)</b>  <i>This isoform has been chosen as the 'canonical' sequence. All positional information in this entry refers to it. This is also the sequence that appears in the downloadable versions of the entry.</i>
<b>Isoform 2 (identifier: P00749-2)</b>  <i>The sequence of this isoform differs from the canonical sequence as follows:</i> 1-29: MRALLARLLLCVLVWSDSKGSNELHQVPS → MVFHLRTRYEQA





Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Featu
Molecule processing					
	Signal peptide	1 – 20	20	 	
	Chain	21 – 431	411	Urokinase-type plasminogen activator	PRO_
	Chain	21 – 177	157	Urokinase-type plasminogen activator long chain A	PRO_
	Chain	156 – 177	22	Urokinase-type plasminogen activator short chain A	PRO_
	Chain	179 – 431	253	Urokinase-type plasminogen activator chain B	PRO_
















## Regions

	Domain	27 – 63	37	EGF-like		
	Domain	70 – 151	82	Kringle		
	Domain	179 – 424	246	Peptidase S1		
	Region	34 – 57	24	Binds urokinase plasminogen activator surface receptor <small>By similarity</small>		
	Region	152 – 177	26	Connecting peptide		

## Sites

	Active site	224	1	Charge relay system		
	Active site	275	1	Charge relay system		
	Active site	376	1	Charge relay system		
	Site	177 – 178	2	Cleavage; during zymogen activation		

## Amino acid modifications

	Modified residue	158	1	Phosphoserine <small>Ref.17</small>		
	Modified residue	323	1	Phosphoserine <small>Ref.17</small>		
	Glycosylation	38	1	O-linked (Fuc) <small>Ref.12</small>		
	Glycosylation	322	1	N-linked (GlcNAc...)		CAR_
	Disulfide bond	31 ↔ 39				
	Disulfide bond	33 ↔ 51				
	Disulfide bond	53 ↔ 62				
	Disulfide bond	70 ↔ 151				
	Disulfide bond	91 ↔ 133				
	Disulfide bond	122 ↔ 146				
	Disulfide bond	168 ↔ 299		Interchain (between A and B chains)		
	Disulfide bond	209 ↔ 225				
	Disulfide bond	217 ↔ 288				
	Disulfide bond	313 ↔ 382				
	Disulfide bond	345 ↔ 361				



<input type="checkbox"/>	Disulfide bond	372 ↔ 400				
<b>Natural variations</b>						
<input checked="" type="checkbox"/>	Alternative sequence	1 – 29	29	MRALL...HQQVPS → MVFHLRTRYEQA in isoform 2.		VSP_
<input type="checkbox"/>	Natural variant	15	1	V → L. [dbSNP:rs2227580] Ref.7		VAR_
<input type="checkbox"/>	Natural variant	141	1	P → L. [dbSNP:rs2227564] Ref.7 Ref.8 Ref.11 Ref.27 Ref.29 Ref.30		VAR_
<input type="checkbox"/>	Natural variant	214	1	I → M		VAR_
<input type="checkbox"/>	Natural variant	231	1	K → Q. [dbSNP:rs2227567] Ref.7		VAR_
<b>Experimental info</b>						
<input type="checkbox"/>	Mutagenesis	158	1	S → E: Abolishes phosphorylation, proadhesive function and ability to induce chemotactic response; when associated with E-323. Ref.17		
<input type="checkbox"/>	Mutagenesis	323	1	S → E: Abolishes phosphorylation, proadhesive function and ability to induce chemotactic response; when associated with E-158. Ref.17		
<input type="checkbox"/>	Sequence conflict	150	1	D → G in BAG60754. Ref.6		
<input type="checkbox"/>	Sequence conflict	151	1	C → W in CAA26535. Ref.2		
<input type="checkbox"/>	Sequence conflict	386	1	G → C in CAA26535. Ref.2		
<input type="checkbox"/>	Sequence conflict	430	1	A → V in CAA26535. Ref.2		

Secondary structure

1 . . . . .

■ Helix

■ Strand

■ Turn

Details...

Sequences

Sequence	Length	Mass (Da)
<div><div><div>■</div></div><div>Isoform 1 [UniParc].</div><div>FASTA</div><div>431</div><div>48,507</div></div> <div>Last modified July 28, 2009. Version 2. Checksum: 62C72400BC23115F</div> <div><div><div>102030405060</div><div>MRALLARLLL CVLVVSDSKG SNELHQVPSN CDCLNGGTCV SNKYFSNIHW CNCPKKFGGQ</div><div>708090100110120</div><div>HCEIDKSKTC YEGNGHFYRG KASTDTMGRP CLPWNSATVL QQTYHAHRSD ALQLGLGKHN</div><div>130140150160170180</div><div>YCRNPDNRRR PWCYVQVGLK PLVQECMVHD CADGKKPSSP PEELKFQCGQ KTLRPRFKII</div><div>190200210220230240</div><div>GGEFTTIENQ PWFAAIYRRH RGGSVTYVCG GSLISPCWVI SATHCFIDYP KKEDIIVYLG</div><div>250260270280290300</div><div>RSRLNSNTQG EMKFEVENLI LHKDYSADTL AHNDIALLK IRSKEGRCAQ PSRTIQTICL</div><div>310320330340350360</div><div>PSMYNDPQFG TSCEITGFGK ENSTDYLYPE QLKMTVVKLI SHRECQPHY YGSEVTTKML</div><div>370380390400410420</div><div>CAADPQWKTD SCQGDGGPL VCSLQGRMTL TGIVSWGRGC ALKDKPGVYT RVSHFLPWIR</div><div>430</div><div>SHTKEENGLA L</div></div></div> <div>« Hide</div>		
<div><div><div>■</div></div><div>Isoform 2.</div><div>FASTA</div><div>414</div><div>46,908</div></div> <div>Checksum: FB35DBE360D7E1CC Show »</div>		

References




« Hide 'large scale' references

[1]

"Cloning and expression of the gene for pro-urokinase in Escherichia coli."  
Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A., Steffens G.J., Heyneker H.L.  
Biotechnology (N.Y.) 3:923-929(1985)  
Cited for: NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).

[2]

"Molecular cloning, sequencing, and expression in Escherichia coli of human preprourokinase cDNA."  
Jacobs P., Cravador A., Loriau R., Brockly F., Colau B., Chuchana P., van Elsen A., Herzog A., Bollen A.  
DNA 4:139-146(1985) [PubMed: 3888571] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).

[3]	<p><b>"Molecular cloning of cDNA coding for human preprourokinase."</b>  Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Arimura H., Nishida M., Suyama T.  Gene 36:183-188(1985) [PubMed: 2415429] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).</p>
[4]	<p><b>"The human urokinase-plasminogen activator gene and its promoter."</b>  Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.  Nucleic Acids Res. 13:2759-2771(1985) [PubMed: 2987867] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA], VARIANT MET-214.</p>
[5]	<p><b>"Cloning of human full-length CDSs in BD Creator(TM) System Donor vector."</b>  Kalnina N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S., Koundinya M., Raphael J.,  Moreira D., Kelley T., LaBaer J., Lin Y., Phelan M., Farmer A.  Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).</p>
[6]	<p><b>"Complete sequencing and characterization of 21,243 full-length human cDNAs."</b>  Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K.,  Sato H., Nagai K., Kimura K., Makita H., Sekine M., Obayashi M., Nishi T., Shibahara T.,  Tanaka T., Ishii S.  Sugano S.  Nat. Genet. 36:40-45(2004) [PubMed: 14702039] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).  <u>Tissue:</u> Mesangial cell.</p>
[7]	<p>SeattleSNPs variation discovery resource  Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA], VARIANTS LEU-15; LEU-141 AND  GLN-231.</p>
[8]	<p><b>"The DNA sequence and comparative analysis of human chromosome 10."</b>  Deloukas P., Earthrowl M.E., Grafham D.V., Rubenfield M., French L., Steward C.A., Sims  S.K., Jones M.C., Searle S., Scott C., Howe K., Hunt S.E., Andrews T.D., Gilbert J.G.R.,  Swarbreck D., Ashurst J.L., Taylor A., Battles J.  Rogers J.  Nature 429:375-381(2004) [PubMed: 15164054] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA], VARIANT LEU-141.</p>
[9]	<p>Mural R.J., Istrail S., Sutton G.G., Florea L., Halpern A.L., Mobarry C.M., Lippert R., Walenz  B., Shatkay H., Dew I., Miller J.R., Flanigan M.J., Edwards N.J., Bolanos R., Fasulo D.,  Halldorsson B.V., Hannenhalli S., Turner R.  Venter J.C.  Submitted (JUL-2005) to the EMBL/GenBank/DDBJ databases  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].</p>
[10]	<p><b>"The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian  Gene Collection (MGC)."</b>  The MGC Project Team  Genome Res. 14:2121-2127(2004) [PubMed: 15489334] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).  <u>Tissue:</u> Lung.</p>
[11]	<p><b>"Characterization of single chain urokinase-type plasminogen activator with a novel  amino-acid substitution in the kringle structure."</b>  Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K., Sawasaki Y.,  Hanada K.  Biochim. Biophys. Acta 1293:83-89(1996) [PubMed: 8652631] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA] OF 21-431, VARIANT LEU-141.</p>
[12]	<p><b>"Characterization of a posttranslational fucosylation in the growth factor domain of  urinary plasminogen activator."</b>  Buko A.M., Kentzer E.J., Petros A., Menon G., Zuiderweg E.R., Sarin V.K.  Proc. Natl. Acad. Sci. U.S.A. 88:3992-3996(1991) [PubMed: 2023947] [Abstract]  <u>Cited for:</u> PROTEIN SEQUENCE OF 21-43, GLYCOSYLATION AT THR-38, MASS  SPECTROMETRY.</p>

[13]	<p><b>"Identification and primary sequence of an unspliced human urokinase poly(A)+ RNA."</b>  Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.  Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984) [PubMed: 6589620] [Abstract]  <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA / MRNA] OF 66-431, VARIANT MET-214.</p>
[14]	<p><b>"The primary structure of high molecular mass urokinase from human urine. The complete amino acid sequence of the A chain."</b>  Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E., Flohe L.  Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982) [PubMed: 6754569] [Abstract]  <u>Cited for:</u> PROTEIN SEQUENCE OF 21-177.</p>
[15]	<p><b>"Human low-molecular-weight urinary urokinase. Partial characterization and preliminary sequence data of the two polypeptide chains."</b>  Schaller J., Nick H., Rickli E.E., Gillesen D., Lergier W., Studer R.O.  Eur. J. Biochem. 125:251-257(1982) [PubMed: 6749491] [Abstract]  <u>Cited for:</u> PROTEIN SEQUENCE OF 156-176 AND 179-224.</p>
[16]	<p><b>"The complete amino acid sequence of low molecular mass urokinase from human urine."</b>  Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.  Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982) [PubMed: 6754572] [Abstract]  <u>Cited for:</u> PROTEIN SEQUENCE OF 158-410.</p>
[17]	<p><b>"Phosphorylation of human pro-urokinase on Ser138/303 impairs its receptor-dependent ability to promote myelomonocytic adherence and motility."</b>  Franco P., Iaccarino C., Chiaradonna F., Brandazza A., Iavarone C., Mastronicola M.R., Nolli M.L., Stoppelli M.P.  J. Cell Biol. 137:779-791(1997) [PubMed: 9151681] [Abstract]  <u>Cited for:</u> PHOSPHORYLATION AT SER-158 AND SER-323, MUTAGENESIS OF SER-158 AND SER-323.</p>
[18]	<p><b>"A urokinase receptor-associated protein with specific collagen binding properties."</b>  Behrendt N., Jensen O.N., Engelholm L.H., Moertz E., Mann M., Danoe K.  J. Biol. Chem. 275:1993-2002(2000) [PubMed: 10636902] [Abstract]  <u>Cited for:</u> INTERACTION WITH MRC2.</p>
[19]	<p><b>"The putative tumor suppressor LRP1B, a novel member of the low density lipoprotein (LDL) receptor family, exhibits both overlapping and distinct properties with the LDL receptor-related protein."</b>  Liu C.-X., Li Y., Obermoeller-McCormick L.M., Schwartz A.L., Bu G.  J. Biol. Chem. 276:28889-28896(2001) [PubMed: 11384978] [Abstract]  <u>Cited for:</u> INTERACTION WITH LRP1B.</p>
[20]	<p><b>"Platelet-derived growth factor D is activated by urokinase plasminogen activator in prostate carcinoma cells."</b>  Ustach C.V., Kim H.-R.C.  Mol. Cell. Biol. 25:6279-6288(2005) [PubMed: 15988036] [Abstract]  <u>Cited for:</u> TISSUE SPECIFICITY.</p>
[21]	<p><b>"Dynamics of the multidomain fibrinolytic protein urokinase from two-dimensional NMR."</b>  Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.  Nature 337:579-582(1989) [PubMed: 2536903] [Abstract]  <u>Cited for:</u> STRUCTURE BY NMR.</p>
[22]	<p><b>"Sequential 1H NMR assignments and secondary structure of the kringle domain from urokinase."</b>  Li X., Smith R.A.G., Dobson C.M.  Biochemistry 31:9562-9571(1992) [PubMed: 1327118] [Abstract]  <u>Cited for:</u> STRUCTURE BY NMR OF 67-155.</p>



[23]	<p><b>"Solution structure of the kringle domain from urokinase-type plasminogen activator."</b>  Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.  J. Mol. Biol. 235:1548-1559(1994) [PubMed: 8107091] [Abstract]  Cited for: STRUCTURE BY NMR OF 67-155.</p>
[24]	<p><b>"The crystal structure of the catalytic domain of human urokinase-type plasminogen activator."</b>  Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D., Dobson C.M., Stuart D.I., Jones E.Y.  Structure 3:681-691(1995) [PubMed: 8591045] [Abstract]  Cited for: X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).</p>
[25]	<p><b>"(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly selective inhibitors of human urokinase."</b>  Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G., Bode W., Magdolen V., Huber R., Moroder L.  Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000) [PubMed: 10805774] [Abstract]  Cited for: X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.</p>
[26]	<p><b>"Structure of human urokinase plasminogen activator in complex with its receptor."</b>  Huai Q., Mazar A.P., Kuo A., Parry G.C., Shaw D.E., Callahan J., Li Y., Yuan C., Bian C., Chen L., Furie B., Furie B.C., Cines D.B., Huang M.  Science 311:656-659(2006) [PubMed: 16456079] [Abstract]  Cited for: X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 31-152 IN COMPLEX WITH PLAUR.</p>
[27]	<p><b>"Detection of polymorphisms in the human urokinase-type plasminogen activator gene."</b>  Conne B., Berczy M., Belin D.  Thromb. Haemost. 77:434-435(1997) [PubMed: 9065988] [Abstract]  Cited for: VARIANT LEU-141.</p>
[28]	<p><b>Erratum</b>  Conne B., Berczy M., Belin D.  Thromb. Haemost. 78:973-973(1997)</p>
[29]	<p><b>"Mutational analysis of the genes encoding urokinase-type plasminogen activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer."</b>  Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W., Creutzburg S., Graeff H., Magdolen V.  Electrophoresis 18:686-689(1997) [PubMed: 9194591] [Abstract]  Cited for: VARIANT LEU-141.</p>
[30]	<p><b>"DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome."</b>  Ley T.J., Mardis E.R., Ding L., Fulton B., McLellan M.D., Chen K., Dooling D., Dunford-Shore B.H., McGrath S., Hickenbotham M., Cook L., Abbott R., Larson D.E., Koboldt D.C., Pohl C., Smith S., Hawkins A., Abbott S., Wilson R.K.  Nature 456:66-72(2008) [PubMed: 18987736] [Abstract]  Cited for: VARIANT [LARGE SCALE ANALYSIS] LEU-141.</p>
+	Additional computationally mapped references.

## Web resources

Wikipedia  
Urokinase entry  
SeattleSNPs

## Cross-references

### Sequence databases

<ul style="list-style-type: none"> <li>⊙ EMBL</li> <li>⊙ GenBank</li> <li>⊙ DDBJ</li> </ul>	M15476 mRNA. Translation: AAA61253.1. X02760 mRNA. Translation: CAA26535.1. D00244 mRNA. Translation: BAA00175.1. K03226 mRNA. Translation: AAC97138.1. X02419 Genomic DNA. Translation: CAA26268.1. AF377330 Genomic DNA. Translation: AAK53822.1. BT007391 mRNA. Translation: AAP36055.1. AK298560 mRNA. Translation: BAG60754.1. AL596247 Genomic DNA. Translation: CAI13969.1. CH471083 Genomic DNA. Translation: EAW54544.1. BC013575 mRNA. Translation: AAH13575.1. D11143 mRNA. Translation: BAA01919.1. K02286 Genomic DNA. Translation: AAA61252.1.
---	--

IPI	IPI00296180. IPI00645018.
-----	------------------------------

PIR	UKHU. A00931.
-----	---------------

RefSeq	NP_001138503.1. NM_001145031.1. NP_002649.1. NM_002658.3.
--------	--

UniGene	Hs.77274.
---------	-----------

### 3D structure databases

<ul style="list-style-type: none"> <li>⊙ PDBe</li> <li>⊙ RCSB PDB</li> <li>⊙ PDBj</li> </ul>	Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
	1C5W	X-ray	1.94	A	156-178	[»]
				B	179-431	[»]
	1C5X	X-ray	1.75	A	156-178	[»]
				B	179-431	[»]
	1C5Y	X-ray	1.65	A	156-178	[»]
				B	179-431	[»]
	1C5Z	X-ray	1.85	A	156-178	[»]
				B	179-431	[»]
	1EJN	X-ray	1.80	A	179-431	[»]
	1F5K	X-ray	1.80	U	179-431	[»]
	1F5L	X-ray	2.10	A	179-431	[»]
	1F92	X-ray	2.60	A	179-431	[»]
	1FV9	X-ray	3.00	A	179-423	[»]
	1GI7	X-ray	1.79	A	156-178	[»]
				B	179-423	[»]
	1GI8	X-ray	1.75	A	156-178	[»]
				B	179-423	[»]
	1GI9	X-ray	1.80	A	156-178	[»]
				B	179-423	[»]
	1GJ7	X-ray	1.50	A	156-178	[»]
				B	179-431	[»]
	1GJ8	X-ray	1.64	A	156-178	[»]
				B	179-431	[»]
	1GJ9	X-ray	1.80	A	156-178	[»]
				B	179-431	[»]
	1GJA	X-ray	1.56	A	156-178	[»]
				B	179-431	[»]
	1GJB	X-ray	1.90	A	156-178	[»]
				B	179-431	[»]
	1GJC	X-ray	1.73	A	156-178	[»]
				B	179-431	[»]
	1GJD	X-ray	1.75	A	156-178	[»]


4/12/2011

STRING	P00749.
<b>Protein family/group databases</b>	
MEROPS	S01.231.
<b>PTM databases</b>	
GlycoSuiteDB	P00749.
PhosphoSite	P00749.
<b>Proteomic databases</b>	
PRIDE	P00749.
<b>Genome annotation databases</b>	
Ensembl	ENST00000372764; ENSP00000361850; ENSG00000122861.
GeneID	5328.
KEGG	hsa:5328.
<b>Organism-specific databases</b>	
CTD	5328.
GeneCards	GC10P069666.
HGNC	HGNC:9052. PLAU.
HPA	HPA008719.
MIM	191840. gene.
neXtProt	NX_P00749.
Orphanet	220436. Quebec platelet disorder.
PharmGKB	PA33382.
GenAtlas	Search...
<b>Phylogenomic databases</b>	
eggNOG	prNOG05724.
HOGENOM	HBG755338.
HOVERGEN	HBG008633.
InParanoid	P00749.
OrthoDB	EOG4229K6.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.21.73. 247.
Pathway_Interaction_DB	amb2_neutrophils_pathway. amb2 Integrin signaling. fgf_pathway. FGF signaling pathway. avb3_opn_pathway. Osteopontin-mediated events.
Reactome	REACT_604. Hemostasis.



<b>Gene expression databases</b>	
ArrayExpress	P00749.
Bgee	P00749.
CleanEx	HS_PLAU.
Genevestigator	P00749.
GermOnline	ENSG00000122861. Homo sapiens.
<b>Family and domain databases</b>	
InterPro	IPR006210. EGF-like. IPR013032. EGF-like_reg_CS. IPR000742. EGF_3. IPR000001. Kringle. IPR013806. Kringle-like. IPR018056. Kringle_CS. IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Gene3D	G3DSA:2.40.20.10. Kringle. 1 hit.
Pfam	PF00051. Kringle. 1 hit. PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN. PR00018. KRINGLE.
SMART	SM00181. EGF. 1 hit. SM00130. KR. 1 hit. SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF57440. Kringle-like. 1 hit. SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS00022. EGF_1. 1 hit. PS01186. EGF_2. False negative. PS50026. EGF_3. 1 hit. PS00021. KRINGLE_1. 1 hit. PS50070. KRINGLE_2. 1 hit. PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
BindingDB	P00749.
DrugBank	DB00594. Amiloride. DB00013. Urokinase.
NextBio	20628.
PMAP-CutDB	P00749.
SOURCE	Search...

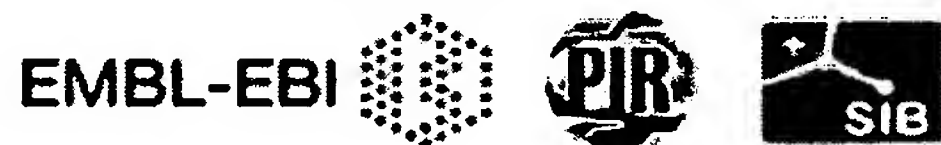
**Entry information**

Entry name	UROK_HUMAN
Accession	Primary (citable) accession number: <b>P00749</b> Secondary accession number(s): B4DPZ2  Q969W6
Entry history	Integrated into July 21, 1986 UniProtKB/Swiss-Prot: Last sequence update: July 28, 2009 Last modified: April 5, 2011 This is version 158 of the entry and version 2 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
Disclaimer	Any medical or genetic information present in this entry is provided for research, educational and informational purposes only. It is not in any way intended to be used as a substitute for professional medical advice, diagnosis, treatment or care.

**Relevant documents**

Human chromosome 10 Human chromosome 10: entries, gene names and cross-references to MIM
Human entries with polymorphisms or disease mutations List of human entries with polymorphisms or disease mutations
Human polymorphisms and disease mutations Index of human polymorphisms and disease mutations
MIM cross-references Online Mendelian Inheritance in Man (MIM) cross-references in UniProtKB/Swiss-Prot
PDB cross-references Index of Protein Data Bank (PDB) cross-references
Peptidase families Classification of peptidase families and list of entries
SIMILARITY comments Index of protein domains and families

© 2002–2011 UniProt Consortium | License &amp; Disclaimer | Contact



## Protein

Translations of Life

Display Settings: GenPept

## pro-urokinase [Homo sapiens]

GenBank: AAA61253.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS AAA61253 431 aa linear PRI 14-JAN-1995  
DEFINITION pro-urokinase [Homo sapiens].  
ACCESSION AAA61253  
VERSION AAA61253.1 GI:340160  
DBSOURCE locus HUMUKPM accession M15476.1  
KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.

REFERENCE 1 (residues 1 to 431)

AUTHORS Holmes,W.E., Pennica,D., Blaber,M., Rey,M.W., Guenzler,W.A.,  
Steffens,G.J. and Heyneker,H.L.

TITLE Cloning and expression of the gene for pro-urokinase in Escherichia  
coli

JOURNAL Biotechnology (N.Y.) 3, 923-929 (1985)

COMMENT Clean copy of sequence [Bio/Technology 3, 923-929 (1985)] kindly  
provided by G.Vehar  
(30-MAY-1986).

Method: conceptual translation.

FEATURES Location/Qualifiers

source 1..431  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="10q24-qter"

Protein 1..431  
/name="pro-urokinase"

sig peptide 1..20  
/gene="PLAU"  
/note="urokinase signal peptide; G00-119-497"

mat peptide 21..431  
/gene="PLAU"  
/product="urokinase; G00-119-497"

Region 67..152  
/region\_name="KR"  
/note="Kringle domain; Kringle domains are believed to  
play a role in binding mediators, such as peptides, other  
proteins, membranes, or phospholipids. They are autonomous  
structural domains, found in a varying number of copies,  
in blood clotting and...; cd00108"  
/db\_xref="CDD:29008"

Site 79  
/site\_type="other"  
/note="putative domain interaction site"  
/db\_xref="CDD:29008"

Site order(94,104,132,134,143)  
/site\_type="other"  
/note="ligand binding site"

Region /db\_xref="CDD:29008"  
179..422  
/region\_name="Tryp\_Spc"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cd00190"  
/db\_xref="CDD:29152"

Site 179  
/site\_type="cleavage"  
/db\_xref="CDD:29152"

Site order(224,275,376)  
/site\_type="active"  
/db\_xref="CDD:29152"

Site order(370,395,397)  
/site\_type="other"  
/note="substrate binding sites"  
/db\_xref="CDD:29152"

CDS 1..431  
/gene="PLAU"  
/coded\_by="M15476.1:77..1372"  
/db\_xref="GDB:G00-119-497"

ORIGIN

1 mrallarlll cvlvvsdskg snelhqvpsn cdclnggtcv snkyfsnihw cncpkkfagg  
61 hceidskctc yegnghfyrg kastdtmgrp clpwnsatvl qqtyahrsd alqlglgkhn  
121 ycrnpgdnrrr pwcyyqvglk plvqecmvhd cadgkpkssp peelkfqcqg ktlrprfkii  
181 ggefttieng pwfaalyrrh rggsavtyvcg gslispcwvi sathcfidyp kkedyivylg  
241 rsrlnsttqg emkfevenli lhdysadt1 ahndiallk irskegrcaq psrtigticl  
301 psmyndpqfg tsceitgfgk enstdyltpe qlkmtvvkli shrecqqphy ygsevttkml  
361 caadpqwktc scqgdsggpl vcslqgrmtl tgivswgrgc alkdpgvyt rvshflpwir  
421 shtkeengla 1

//

## Protein

Translations of Life

Display Settings: GenPept

## unnamed protein product [Homo sapiens]

GenBank: CAA26535.1

[FASTA](#) [Graphics](#)[Go to:](#)

LOCUS CAA26535 431 aa linear PRI 30-MAR-1995  
DEFINITION unnamed protein product [Homo sapiens].  
ACCESSION CAA26535  
VERSION CAA26535.1 GI:35298  
DBSOURCE embl accession X02760.1  
KEYWORDS .  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 431)  
AUTHORS Jacobs,P., Cravador,A., Loriau,R., Brockly,F., Colau,B.,  
Chuchana,P., van Elsen,A., Herzog,A. and Bollen,A.  
TITLE Molecular cloning, sequencing, and expression in Escherichia coli  
of human preprourokinase cDNA  
JOURNAL DNA 4 (2), 139-146 (1985)  
PUBMED 3888571  
COMMENT Data kindly reviewed (5-DEC-1985) by A.Bollen.  
FEATURES  
Location/Qualifiers  
source 1..431  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
Protein 1..431  
/name="unnamed protein product"  
sig\_peptide 1..20  
/note="signal peptide (aa -20 to -1)"  
Region 67..149  
/region\_name="KR"  
/note="Kringle domain; Kringle domains are believed to  
play a role in binding mediators, such as peptides, other  
proteins, membranes, or phospholipids. They are autonomous  
structural domains, found in a varying number of copies,  
in blood clotting and...; cd00108"  
/db\_xref="CDD:29008"  
Site 79  
/site\_type="other"  
/note="putative domain interaction site"  
/db\_xref="CDD:29008"  
Site order(94,104,132,134,143)  
/site\_type="other"  
/note="ligand binding site"  
/db\_xref="CDD:29008"  
Region 179..422  
/region\_name="Tryp\_SpC"  
/note="Trypsin-like serine protease; Many of these are  
synthesized as inactive precursor zymogens that are  
cleaved during limited proteolysis to generate their  
active forms. Alignment contains also inactive enzymes  
that have substitutions of the catalytic...; cd00190"  
/db\_xref="CDD:29152"  
Site 179  
/site\_type="cleavage"  
/db\_xref="CDD:29152"  
Site order(224,275,376)  
/site\_type="active"  
/db\_xref="CDD:29152"  
Site order(370,395,397)  
/site\_type="other"  
/note="substrate binding sites"  
/db\_xref="CDD:29152"  
CDS 1..431  
/coded\_by="X02760.1:41..1336"  
/note="urokinase precursor"  
/db\_xref="GDB:119497"  
/db\_xref="GOA:P00749"  
/db\_xref="HGNC:9052"  
/db\_xref="InterPro:IPR000001"  
/db\_xref="InterPro:IPR000742"  
/db\_xref="InterPro:IPR001254"  
/db\_xref="InterPro:IPR001314"  
/db\_xref="InterPro:IPR006210"  
/db\_xref="InterPro:IPR009003"  
/db\_xref="InterPro:IPR013032"  
/db\_xref="InterPro:IPR013806"  
/db\_xref="InterPro:IPR018056"  
/db\_xref="InterPro:IPR018114"  
/db\_xref="PDB:1C5W"  
/db\_xref="PDB:1C5X"  
/db\_xref="PDB:1C5Y"  
/db\_xref="PDB:1C5Z"  
/db\_xref="PDB:1EJN"  
/db\_xref="PDB:1F5K"  
/db\_xref="PDB:1F5L"  
/db\_xref="PDB:1F92"  
/db\_xref="PDB:1FV9"  
/db\_xref="PDB:1GI7"  
/db\_xref="PDB:1GI8"



/db\_xref="PDB:1GI9"  
/db\_xref="PDB:1GJ7"  
/db\_xref="PDB:1GJ8"  
/db\_xref="PDB:1GJ9"  
/db\_xref="PDB:1GJA"  
/db\_xref="PDB:1GJB"  
/db\_xref="PDB:1GJC"  
/db\_xref="PDB:1GJD"  
/db\_xref="PDB:1KDU"  
/db\_xref="PDB:1LMW"  
/db\_xref="PDB:1O3P"  
/db\_xref="PDB:1O5A"  
/db\_xref="PDB:1O5B"  
/db\_xref="PDB:1O5C"  
/db\_xref="PDB:1OWD"  
/db\_xref="PDB:1OWE"  
/db\_xref="PDB:1OWH"  
/db\_xref="PDB:1OWI"  
/db\_xref="PDB:1OWJ"  
/db\_xref="PDB:1OWK"  
/db\_xref="PDB:1SC8"  
/db\_xref="PDB:1SQA"  
/db\_xref="PDB:1SQO"  
/db\_xref="PDB:1SQT"  
/db\_xref="PDB:1U6Q"  
/db\_xref="PDB:1URK"  
/db\_xref="PDB:1VJ9"  
/db\_xref="PDB:1VJA"  
/db\_xref="PDB:1W0Z"  
/db\_xref="PDB:1W10"  
/db\_xref="PDB:1W11"  
/db\_xref="PDB:1W12"  
/db\_xref="PDB:1W13"  
/db\_xref="PDB:1W14"  
/db\_xref="PDB:2FD6"  
/db\_xref="PDB:2I9A"  
/db\_xref="PDB:2I9B"  
/db\_xref="PDB:2NWN"  
/db\_xref="PDB:2O8T"  
/db\_xref="PDB:2O8U"  
/db\_xref="PDB:2O8W"  
/db\_xref="PDB:2R2W"  
/db\_xref="PDB:2VIN"  
/db\_xref="PDB:2VIO"  
/db\_xref="PDB:2VIP"  
/db\_xref="PDB:2VIQ"  
/db\_xref="PDB:2VIV"  
/db\_xref="PDB:2VIW"  
/db\_xref="PDB:2VNT"  
/db\_xref="PDB:3BT1"  
/db\_xref="PDB:3BT2"  
/db\_xref="PDB:3IG6"  
/db\_xref="PDB:3KGP"  
/db\_xref="PDB:3KHV"  
/db\_xref="PDB:3KID"  
/db\_xref="PDB:3M61"  
/db\_xref="PDB:3MHW"  
/db\_xref="UniProtKB/Swiss-Prot:P00749"

## ORIGIN

```
1  mrallarlll  cvlvvsdskg  snelhqvpn  cdclnggtcv  snkyfsnihw  cncpkkfagg
61 hceidkskctc yegnghfyrg kastdtmgrp clpwnsatvl qqtyhahrsd alqlglgkhn
121 ycrnpdnrrr pwcyvqvglk plvqecmvhd wadgkpkssp peelkfqcqg ktlrprfkii
181 ggefttieng pwfaaiyrrh rggsvttyvcg gslispcwvi sathcfidyp kkedyivylg
241 rsrlnsntqg emkfevenli lhkdysadt1 ahndiallk irskegrcaq partiqticl
301 psmyndpqfg tsceitgfgk enstdylpe qlkmtvvkli shrecqgphy ygsevttkml
361 caadpqwktd scqgdsggpl vcslqcrmtl tgivswgrgc alkdkgvyt rvshflpwir
421 shtkeenglv 1
```

//

# Protein

Translations of Life

Display Settings: GenPept

## preprourokinase [Homo sapiens]

GenBank: AAC97138.1

FASTA Graphics

Go to:

```

LOCUS      AAC97138          431 aa          linear   PRI 18-DEC-1998
DEFINITION preprourokinase [Homo sapiens].
ACCESSION  AAC97138
VERSION    AAC97138.1  GI:340158
DBSOURCE   locus HUMUKM1 accession K03226.1
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1 (residues 1 to 431)
  AUTHORS   Nagai,M., Hiramatsu,R., Kaneda,T., Hayasuke,N., Arimura,H.,
            Nishida,M. and Suyama,T.
  TITLE     Molecular cloning of cDNA coding for human preprourokinase
  JOURNAL   Gene 36 (1-2), 183-188 (1985)
  PUBMED    2415429
COMMENT     Method: conceptual translation.
FEATURES             Location/Qualifiers
     source          1..431
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /map="10q24-qter"
     Protein         1..431
                     /name="preprourokinase"
     sig_peptide     1..20
     mat_peptide     21..177
                     /product="urokinase A chain"
     Region          67..152
                     /region_name="KR"
                     /note="Kringle domain; Kringle domains are believed to
                     play a role in binding mediators, such as peptides, other
                     proteins, membranes, or phospholipids. They are autonomous
                     structural domains, found in a varying number of copies,
                     in blood clotting and...; cd00108"
                     /db_xref="CDD:29008"
     Site            79
                     /site_type="other"
                     /note="putative domain interaction site"
                     /db_xref="CDD:29008"
     Site            order(94,104,132,134,143)
                     /site_type="other"
                     /note="ligand binding site"
                     /db_xref="CDD:29008"
     mat_peptide     156..176
                     /product="urokinase A1 chain"
     mat_peptide     179..431
                     /product="urokinase B chain"
     Region          179..422
                     /region_name="Tryp_SPC"
                     /note="Trypsin-like serine protease; Many of these are
                     synthesized as inactive precursor zymogens that are
                     cleaved during limited proteolysis to generate their
                     active forms. Alignment contains also inactive enzymes
                     that have substitutions of the catalytic...; cd00190"
                     /db_xref="CDD:29152"
     Site            179
                     /site_type="cleavage"
                     /db_xref="CDD:29152"
     Site            order(224,275,376)
                     /site_type="active"
                     /db_xref="CDD:29152"
     Site            order(370,395,397)
                     /site_type="other"
                     /note="substrate binding sites"
                     /db_xref="CDD:29152"
     CDS             1..431
                     /gene="PLAU"
                     /coded_by="K03226.1:81..1376"
                     /db_xref="GDB:G00-119-497"

```

```

ORIGIN
1  mrallarlll cvlvvsdskg snelhqvpsn cdclnggtcv snkyfsnihw cncpkkfagg
61 hceidkskctc yegnghfyrg kastdtmgrp clpwnsatvl qqtyhahrsd alqlglgkhn
121 ycrnpdnrrr pwcyyvgvlg plvqecmvhd cadgkpkssp peelkfqcgg ktlrprfkii
181 ggefttieng pwfaaiyrrh rggsvtyvcg glispcwvi sathcfidyp kkedyivylg
241 rsrlnsntgg emkfevenli lhdysadtl ahndiallk irskegrcaq psrtiqticl
301 psmyndpqfg tscetgfgk enstdylpe qlkntvvkli shrecqgphy ygsevttkml
361 caadpqwktd scqgdsaggpl vcslggrmtl tgivswgrgc alkdkgpyvt rvshflpwir
421 shtkeengla 1
//

```

**P06869 (UROK\_MOUSE) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 114.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Urokinase-type plasminogen activator</b> Short name=U-plasminogen activator Short name=uPA EC=3.4.21.73 <i>Cleaved into the following 3 chains:</i> 1. <b>Urokinase-type plasminogen activator long chain A</b> 2. <b>Urokinase-type plasminogen activator short chain A</b> 3. <b>Urokinase-type plasminogen activator chain B</b>
Gene names	Name:Plau
Organism	<b>Mus musculus (Mouse)</b>
Taxonomic identifier	10090 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires › Rodentia › Sciurognathi › Muroidea › Muridae › Murinae › Mus › Mus

**Protein attributes**

Sequence length	433 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Specifically cleave the zymogen plasminogen to form the active enzyme plasmin.
Catalytic activity	Specific cleavage of Arg- -Val bond in plasminogen to form plasmin.
Subunit structure	Found in high and low molecular mass forms. Each consists of two chains, A and B. The high molecular mass form contains a long chain A which is cleaved to yield a short chain A. Binds LRP1B; binding is followed by internalization and degradation. Interacts with MRC2. Interacts with PLAUR <a href="#">By similarity</a> .
Subcellular location	Secreted.
Sequence similarities	Belongs to the peptidase S1 family. Contains 1 EGF-like domain. Contains 1 kringle domain.

		Contains 1 peptidase S1 domain.				
Ontologies						
Keywords						
Biological process		Plasminogen activation				
Cellular component		Secreted				
Domain		EGF-like domain Kringle Signal				
Molecular function		Hydrolase Protease Serine protease				
PTM		Disulfide bond Zymogen				
Technical term		3D-structure				
Gene Ontology (GO)						
Biological process		proteolysis Inferred from electronic annotation. Source: InterPro  regulation of cell proliferation Inferred from genetic interaction. Source: MGI  response to hypoxia Inferred from mutant phenotype. Source: MGI  smooth muscle cell migration Inferred from mutant phenotype. Source: MGI				
Cellular component		extracellular region Inferred from electronic annotation. Source: UniProtKB-SubCell				
Molecular function		serine-type endopeptidase activity Inferred from electronic annotation. Source: UniProtKB-KW				
Complete GO annotation...						
Sequence annotation (Features)						
	Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
Molecule processing						
	Signal peptide	1 – 20	20	Potential		
	Chain	21 – 433	413	Urokinase-type plasminogen activator		PRO_0000028322
	Chain	21 – 178	158	Urokinase-type		PRO_0000028323



<input checked="" type="checkbox"/>	Chain	157 – 178	22	plasminogen activator long chain A <small>By similarity</small>		PRO_0000028324
<input checked="" type="checkbox"/>	Chain	180 – 433	254	Urokinase-type plasminogen activator short chain A <small>By similarity</small>		PRO_0000028325
<b>Regions</b>						
<input checked="" type="checkbox"/>	Domain	28 – 64	37	EGF-like		
<input checked="" type="checkbox"/>	Domain	71 – 152	82	Kringle		
<input checked="" type="checkbox"/>	Domain	180 – 426	247	Peptidase S1		
<input checked="" type="checkbox"/>	Region	35 – 58	24	Binds urokinase plasminogen activator surface receptor <small>By similarity</small>		
<input checked="" type="checkbox"/>	Region	153 – 179	27	Connecting peptide		
<b>Sites</b>						
<input type="checkbox"/>	Active site	226	1	Charge relay system		
<input type="checkbox"/>	Active site	277	1	Charge relay system		
<input type="checkbox"/>	Active site	378	1	Charge relay system		
<b>Amino acid modifications</b>						
<input type="checkbox"/>	Disulfide bond	32 ↔ 40		<small>Ref.3</small>		
<input type="checkbox"/>	Disulfide bond	34 ↔ 52		<small>Ref.3</small>		
<input type="checkbox"/>	Disulfide bond	54 ↔ 63		<small>Ref.3</small>		
<input type="checkbox"/>	Disulfide bond	71 ↔ 152		<small>Ref.3</small>		
<input type="checkbox"/>	Disulfide bond	92 ↔ 134		<small>Ref.3</small>		
<input type="checkbox"/>	Disulfide bond	123 ↔ 147		<small>Ref.3</small>		

<input type="checkbox"/>	Disulfide bond	169 ↔ 301		Interchain (between A and B chains) <div>By similarity</div>	<div></div> <div></div>	
<input type="checkbox"/>	Disulfide bond	211 ↔ 227		<div>By similarity</div>	<div></div> <div></div>	
<input type="checkbox"/>	Disulfide bond	219 ↔ 290		<div>By similarity</div>	<div></div> <div></div>	
<input type="checkbox"/>	Disulfide bond	315 ↔ 384		<div>By similarity</div>	<div></div> <div></div>	
<input type="checkbox"/>	Disulfide bond	347 ↔ 363		<div>By similarity</div>	<div></div> <div></div>	
<input type="checkbox"/>	Disulfide bond	374 ↔ 402		<div>By similarity</div>	<div></div> <div></div>	

Secondary structure

1. .... 433

■ Helix

■ Strand

■ Turn

Details...

Sequences

Sequence

Length

Mass (Da)

☐

P06869 [UniParc].

FASTA

433

48,268

Last modified January 1, 1988. Version 1.

Checksum: A99C35F6250443F9

10202030405060

MKVWLASLFLCALVVKNSEGGSVLGAPDES

708090100110120

EHCEIDASKTCYHGNGDSYRGKANTDTKGR

130140150160170180

NYCRNPDNQKRPWCYVQIGLRQFVQECMVH

190200210220230240

VGGEFTEVENQPWFAAIYQKNKGGSPPSFK

250260270280290300

LGQSKESSYNPGEMKFEVEQLILHEYYRED

310320330340350360

CLPPRFTDAPFGSDCEITGF

370380390400410420

MLCAADPEWKTDSCCKGDSGG

430

IQSHIGEELG LAF

« Hide

References

[1] "Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase-type plasminogen activator."

Belin D., Vassalli J.-D., Combepine C., Godeau F., Nagamine Y., Reich E., Kocher H.P., Duvoisin R.M.

	Eur. J. Biochem. 148:225-232(1985) [PubMed: 2985383] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].
[2]	<b>"The murine urokinase-type plasminogen activator gene."</b> Degen S.J.F., Heckel J.L., Reich E., Degen J.L. Biochemistry 26:8270-8279(1987) [PubMed: 2831940] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA].
[3]	<b>"Structure-based engineering of species selectivity in the interaction between urokinase and its receptor: implication for preclinical cancer therapy."</b> Lin L., Gardsvoll H., Huai Q., Huang M., Ploug M. J. Biol. Chem. 285:10982-10992(2010) [PubMed: 20133942] [Abstract] <u>Cited for:</u> X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS) OF 21-154 IN COMPLEX WITH PLAUR, DISULFIDE BONDS.
+	Additional computationally mapped references.

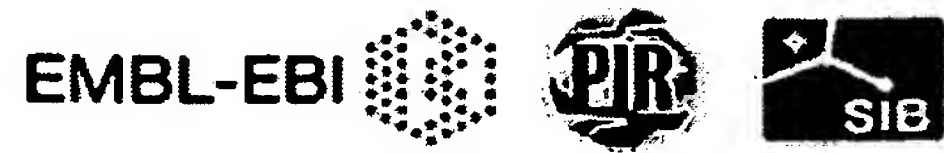
Cross-references

Sequence databases													
<div><div><div></div><div>EMBL</div></div><div><div></div><div>GenBank</div></div><div><div></div><div>DDBJ</div></div></div>	X02389 mRNA. Translation: CAA26231.1. M17922 Genomic DNA. Translation: AAA40539.1.												
IPI	IPI00129102.												
PIR	UKMS. A29420.												
RefSeq	NP_032899.1. NM_008873.3.												
UniGene	Mm.4183.												
3D structure databases													
<div><div><div></div><div>PDBe</div></div><div><div></div><div>RCSB PDB</div></div><div><div></div><div>PDBj</div></div></div>	<table><tr><th>Entry</th><th>Method</th><th>Resolution (Å)</th><th>Chain</th><th>Positions</th><th>PDBsum</th></tr><tr><td>3LAQ</td><td>X-ray</td><td>3.20</td><td>A/B</td><td>21-154</td><td>[»]</td></tr></table>	Entry	Method	Resolution (Å)	Chain	Positions	PDBsum	3LAQ	X-ray	3.20	A/B	21-154	[»]
Entry	Method	Resolution (Å)	Chain	Positions	PDBsum								
3LAQ	X-ray	3.20	A/B	21-154	[»]								
ProteinModelPortal	P06869.												
SMR	P06869. Positions 28-152, 180-424.												
ModBase	Search...												
Protein-protein interaction databases													
STRING	P06869.												
Protein family/group databases													
MEROPS	S01.231.												
PTM databases													
PhosphoSite	P06869.												
Proteomic databases													
PRIDE	P06869.												

<b>Genome annotation databases</b>	
Ensembl	ENSMUST00000022368; ENSMUSP00000022368; ENSMUSG00000021822.
GeneID	18792.
KEGG	mmu:18792.
UCSC	uc007skx.1. mouse.
<b>Organism-specific databases</b>	
CTD	18792.
MGI	MGI:97611. Plau.
<b>Phylogenomic databases</b>	
eggNOG	roNOG07876.
HOGENOM	HBG755338.
HOVERGEN	HBG008633.
InParanoid	P06869.
OMA	QPWFAAI.
OrthoDB	EOG4229K6.
PhylomeDB	P06869.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.21.73. 244.
<b>Gene expression databases</b>	
ArrayExpress	P06869.
Bgee	P06869.
CleanEx	MM_PLAU.
Genevestigator	P06869.
GermOnline	ENSMUSG00000021822. Mus musculus.
<b>Family and domain databases</b>	
InterPro	IPR006210. EGF-like. IPR013032. EGF-like_reg_CS. IPR000742. EGF_3. IPR000001. Kringle. IPR013806. Kringle-like. IPR018056. Kringle_CS. IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Gene3D	G3DSA:2.40.20.10. Kringle. 1 hit.
Pfam	PF00051. Kringle. 1 hit. PF00089. Trypsin. 1 hit. [Graphical view]






PRINTS	PR00722. CHYMOTRYPSIN. PR00018. KRINGLE.
SMART	SM00181. EGF. 1 hit. SM00130. KR. 1 hit. SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF57440. Kringle-like. 1 hit. SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS00022. EGF_1. 1 hit. PS01186. EGF_2. False negative. PS50026. EGF_3. 1 hit. PS00021. KRINGLE_1. 1 hit. PS50070. KRINGLE_2. 1 hit. PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. False negative. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Other Resources</b>	
NextBio	295072.
SOURCE	Search...
<b>Entry information</b>	
Entry name	UROK_MOUSE
Accession	Primary (citable) accession number: <b>P06869</b>
Entry history	Integrated into      January 1, 1988 UniProtKB/Swiss- Prot: Last sequence      January 1, 1988 update: Last modified:      April 5, 2011 This is version 114 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Relevant documents</b>	
MGD cross-references Mouse Genome Database (MGD) cross-references in UniProtKB/Swiss-Prot	
PDB cross-references Index of Protein Data Bank (PDB) cross-references	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	



**Q8MHY7 (UROK\_RABIT)** ★ Reviewed, UniProtKB/Swiss-Prot

Last modified April 5, 2011. Version 54.

Names and origin	
Protein names	<p><i>Recommended name:</i> <b>Urokinase-type plasminogen activator</b> Short name=U-plasminogen activator Short name=uPA EC=3.4.21.73</p> <p><u>Cleaved into the following 3 chains:</u></p> <ol style="list-style-type: none"><li>1. <b><i>Urokinase-type plasminogen activator long chain</i></b> <b>A</b></li><li>2. <b><i>Urokinase-type plasminogen activator short chain</i></b> <b>A</b></li><li>3. <b><i>Urokinase-type plasminogen activator chain B</i></b></li></ol>
Gene names	Name:PLAU
Organism	<i>Oryctolagus cuniculus</i> (Rabbit)
Taxonomic identifier	9986 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires › Lagomorpha › Leporidae › <i>Oryctolagus</i>
Protein attributes	
Sequence length	433 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at transcript level.
General annotation (Comments)	
Function	Specifically cleave the zymogen plasminogen to form the active enzyme plasmin.
Catalytic activity	Specific cleavage of Arg- -Val bond in plasminogen to form plasmin.
Subunit structure	Found in high and low molecular mass forms. Each consists of two chains, A and B. The high molecular mass form contains a long chain A which is cleaved to yield a short chain A. Binds LRP1B; binding is followed by internalization and degradation. Interacts with MRC2. Interacts with PLAUR <a href="#">[By similarity]</a> .
Subcellular location	Secreted <a href="#">[By similarity]</a> .
Post-translational modification	Phosphorylation of Ser-325 abolishes proadhesive ability but does not interfere with receptor binding <a href="#">[By similarity]</a> .
Sequence similarities	Belongs to the peptidase S1 family. Contains 1 EGF-like domain.

Contains 1 kringle domain.					
Contains 1 peptidase S1 domain.					
Ontologies					
Keywords					
Biological process	Plasminogen activation				
Cellular component	Secreted				
Domain	EGF-like domain Kringle Signal				
Molecular function	Hydrolase Protease Serine protease				
PTM	Disulfide bond Glycoprotein Phosphoprotein Zymogen				
Gene Ontology (GO)					
Biological process	proteolysis Inferred from electronic annotation. Source: InterPro				
Cellular component	extracellular region Inferred from electronic annotation. Source: UniProtKB-SubCell				
Molecular function	serine-type endopeptidase activity Inferred from electronic annotation. Source: UniProtKB-KW				
Complete GO annotation...					
Sequence annotation (Features)					
Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
Molecule processing					
	Signal peptide	1 – 16	16	<div>Potential</div>	
	Chain	17 – 433	417	Urokinase-type plasminogen activator	PRO_000028
	Chain	21 – 179	159	Urokinase-type plasminogen activator long chain A <div>By similarity</div>	PRO_000028



<input checked="" type="checkbox"/>	Chain	157 – 179	23	Urokinase-type plasminogen activator short chain A <small>By similarity</small>		PRO_000028
<input checked="" type="checkbox"/>	Chain	181 – 433	253	Urokinase-type plasminogen activator chain B <small>By similarity</small>		PRO_000028

**Regions**

<input checked="" type="checkbox"/>	Domain	29 – 65	37	EGF-like		
<input checked="" type="checkbox"/>	Domain	71 – 153	83	Kringle		
<input checked="" type="checkbox"/>	Domain	181 – 426	246	Peptidase S1		
<input checked="" type="checkbox"/>	Region	36 – 59	24	Binds urokinase plasminogen activator surface receptor <small>By similarity</small>		
<input checked="" type="checkbox"/>	Region	154 – 180	27	Connecting peptide <small>By similarity</small>		

**Sites**

<input type="checkbox"/>	Active site	226	1	Charge relay system <small>By similarity</small>		
<input type="checkbox"/>	Active site	277	1	Charge relay system <small>By similarity</small>		
<input type="checkbox"/>	Active site	378	1	Charge relay system <small>By similarity</small>		

**Amino acid modifications**

<input type="checkbox"/>	Modified residue	325	1	Phosphoserine <small>By similarity</small>		
<input type="checkbox"/>	Glycosylation	324	1	N-linked (GlcNAc...) <small>Potential</small>		
<input type="checkbox"/>	Disulfide bond	33 ↔ 41		<small>By similarity</small>		
<input type="checkbox"/>	Disulfide bond	35 ↔ 53		<small>By similarity</small>		
<input type="checkbox"/>	Disulfide bond	55 ↔ 64		<small>By similarity</small>		
<input type="checkbox"/>	Disulfide bond	72 ↔ 153		<small>By similarity</small>		

<input type="checkbox"/>	Disulfide bond	93 ↔ 135		By similarity	I I	
<input type="checkbox"/>	Disulfide bond	124 ↔ 148		By similarity	I I	
<input type="checkbox"/>	Disulfide bond	170 ↔ 301		Interchain (between A and B chains) By similarity	I I	
<input type="checkbox"/>	Disulfide bond	211 ↔ 227		By similarity	I I	
<input type="checkbox"/>	Disulfide bond	219 ↔ 290		By similarity	I I	
<input type="checkbox"/>	Disulfide bond	315 ↔ 384		By similarity	I I	
<input type="checkbox"/>	Disulfide bond	347 ↔ 363		By similarity	I I	
<input type="checkbox"/>	Disulfide bond	374 ↔ 402		By similarity	I I	

Experimental info

<input type="checkbox"/>	Sequence conflict	268	1	R → S in AAM83187. Ref.2	I	
--------------------------	-------------------	-----	---	--------------------------------	---	--

Sequences

Sequence	Length	Mass (Da)
<input type="checkbox"/> Q8MHY7 [UniParc]. Last modified October 1, 2002. Version 1. Checksum: 6DD35A371010A6EE	FASTA 433	48,444
<div><div><div><div>102030405060</div><div>MRVLLVCLLLCALVVS DSEGSHELHGVSDASNCGCLNGGT CVTYKYFSNIWRCNCPKKFQ</div></div><div><div>708090100110120</div><div>GEHCEIDTLKTCYHGDGHSYRGKANTDIMDRPCLAWSANVLTkTYHAHRPDALQLGLGK</div></div><div><div>130140150160170180</div><div>HNYCRNPdHQRRPWCYVQVG LKQLIQECKVHDCSSGKKPALPPGKLEFQC GQKALRPRFK</div></div><div><div>190200210220230240</div><div>IIGGEFTIIE NQPWFAAIYRHRGGSVTYVCGGSLISPCWVVSATHCFIN HQKKEDYIVY</div></div><div><div>250260270280290300</div><div>LGRSRLNSMT PGEMKFEVEQLILHEGYRAD TLAHHNDIAL LKILSNNGQC AQP SRSIQT I</div></div><div><div>310320330340350360</div><div>CLPPWNADPN FGTSCEITGF GKENSTDYLY PEQLKMTVVKLVS YQECQP HYYGSEVTTK</div></div><div><div>370380390400410420</div><div>MLCAADPQWE TDSCQGDSSGPLVCSVQGRMTLTGIVSWGR GCALKNKPGV YTRVSRFLPW</div></div><div><div>430</div><div>IRSHIGEENG LAL</div></div></div><div>« Hide</div></div>		

References

[1]

"Downregulation of urokinase-type and tissue-type plasminogen activators in a rabbit model of renal ischemia/reperfusion."  
Sugiki M., Omura S., Yoshida E., Itoh H., Kataoka H., Maruyama M.

	J. Biochem. 132:501-508(2002) [PubMed: 12204121] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].
[2]	<b>"Increased expression of urokinase during atherosclerotic lesion development causes arterial constriction and lumen loss, and accelerates lesion growth."</b> Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R., Dichek D.A. Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002) [PubMed: 12149463] [Abstract] <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].
[3]	<b>"<i>Oryctolagus cuniculus</i> urokinase-type plasminogen activator, mRNA, complete cds."</b> Yano W., Watanabe M. Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].
[4]	<b>"Rabbit urokinase plasminogen activator (genomic DNA)."</b> Graf S.K., Wen S., Dichek D.A. Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA].

## Cross-references

### Sequence databases

⊕ EMBL	AY029517 mRNA. Translation: AAK40239.1.
⊕ GenBank	AY122285 mRNA. Translation: AAM83187.1.
⊕ DDBJ	AB087224 mRNA. Translation: BAC02685.1.
	AY453856 Genomic DNA. Translation: AAR21232.1.

RefSeq	NP_001075480.1. NM_001082011.1.
--------	---------------------------------

UniGene	Ocu.2313.
---------	-----------

### 3D structure databases

HSSP	HSSP built from PDB template 1LMW based on UniProtKB P00749.
------	--

ProteinModelPortal	Q8MHY7.
--------------------	---------

SMR	Q8MHY7. Positions 31-157, 168-428.
-----	------------------------------------

ModBase	Search...
---------	-----------

### Protein family/group databases

MEROPS	S01.231.
--------	----------

### Genome annotation databases

Ensembl	ENSOCUT00000029709; ENSOCUP00000023296; ENSOCUG00000003155.
---------	--

GeneID	100008633.
--------	------------

### Organism-specific databases

CTD	100008633.
-----	------------

### Phylogenomic databases

HOVERGEN	HBG008633.
----------	------------

OrthoDB	EOG4229K6.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.21.73. 255.
<b>Family and domain databases</b>	
InterPro	IPR006210. EGF-like. IPR013032. EGF-like_reg_CS. IPR000742. EGF_3. IPR000001. Kringle. IPR013806. Kringle-like. IPR018056. Kringle_CS. IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Gene3D	G3DSA:2.40.20.10. Kringle. 1 hit.
Pfam	PF00051. Kringle. 1 hit. PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN. PR00018. KRINGLE.
SMART	SM00181. EGF. 1 hit. SM00130. KR. 1 hit. SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF57440. Kringle-like. 1 hit. SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS00022. EGF_1. 1 hit. PS01186. EGF_2. False negative. PS50026. EGF_3. 1 hit. PS00021. KRINGLE_1. 1 hit. PS50070. KRINGLE_2. 1 hit. PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Entry information</b>	
Entry name	UROK_RABIT
Accession	Primary (citable) accession number: <b>Q8MHY7</b> Secondary accession number(s): Q8MILO
Entry history	Integrated into      May 1, 2007 UniProtKB/Swiss- Prot: Last sequence      October 1, 2002 update: Last modified:      April 5, 2011 This is version 54 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)



Annotation program

Chordata Protein Annotation Program

**Relevant documents**

---

Peptidase families

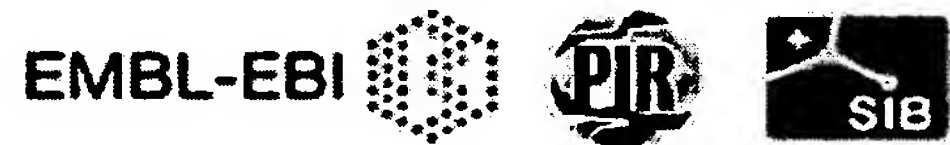
Classification of peptidase families and list of entries

---

SIMILARITY comments

Index of protein domains and families

© 2002–2011 UniProt Consortium | License &amp; Disclaimer | Contact



**P29598 (UROK\_RAT) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 103.

**Names and origin**




Protein names	<i>Recommended name:</i> <b>Urokinase-type plasminogen activator</b> Short name=U-plasminogen activator Short name=uPA EC=3.4.21.73 <i>Cleaved into the following 3 chains:</i> 1. <b>Urokinase-type plasminogen activator long chain A</b> 2. <b>Urokinase-type plasminogen activator short chain A</b> 3. <b>Urokinase-type plasminogen activator chain B</b>
Gene names	Name:Plau
Organism	<b>Rattus norvegicus (Rat)</b>
Taxonomic identifier	10116 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Glires › Rodentia › Sciurognathi › Muroidea › Muridae › Murinae › Rattus

**Protein attributes**

Sequence length	432 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

**General annotation (Comments)**

Function	Specifically cleave the zymogen plasminogen to form the active enzyme plasmin.
Catalytic activity	Specific cleavage of Arg- -Val bond in plasminogen to form plasmin.
Subunit structure	Found in high and low molecular mass forms. Each consists of two chains, A and B. The high molecular mass form contains a long chain A which is cleaved to yield a short chain A. Binds LRP1B; binding is followed by internalization and degradation. Interacts with MRC2. Interacts with PLAUR <span>By similarity</span> .
Subcellular location	Secreted <span>By similarity</span> .
Sequence similarities	Belongs to the peptidase S1 family. Contains 1 EGF-like domain. Contains 1 kringle domain.

		Contains 1 peptidase S1 domain.																									
Ontologies																											
Keywords																											
Biological process		Plasminogen activation																									
Cellular component		Secreted																									
Domain		EGF-like domain Kringle Signal																									
Molecular function		Hydrolase Protease Serine protease																									
PTM		Disulfide bond Zymogen																									
Gene Ontology (GO)																											
Biological process		angiogenesis Inferred from mutant phenotype. Source: RGD chemotaxis Traceable author statement. Source: RGD embryo implantation Inferred from expression pattern. Source: RGD proteolysis Traceable author statement. Source: RGD response to hyperoxia Inferred from expression pattern. Source: RGD signal transduction Traceable author statement. Source: RGD skeletal muscle tissue regeneration Inferred from mutant phenotype. Source: RGD																									
Cellular component		extracellular region Inferred from electronic annotation. Source: UniProtKB-SubCell																									
Molecular function		serine-type endopeptidase activity Inferred from electronic annotation. Source: UniProtKB-KW																									
Complete GO annotation...																											
Sequence annotation (Features)																											
<table><tr><td>Feature key</td><td>Position(s)</td><td>Length</td><td>Description</td><td>Graphical view</td><td colspan="2">Feature identifier</td></tr><tr><td colspan="7">Molecule processing</td></tr><tr><td></td><td>Signal peptide</td><td>1 – 19</td><td>19</td><td>Potential</td><td colspan="2"></td></tr></table>							Feature key	Position(s)	Length	Description	Graphical view	Feature identifier		Molecule processing								Signal peptide	1 – 19	19	Potential		
Feature key	Position(s)	Length	Description	Graphical view	Feature identifier																						
Molecule processing																											
	Signal peptide	1 – 19	19	Potential																							

<input checked="" type="checkbox"/>	Chain	20 – 432	413	Urokinase-type plasminogen activator		PRO_00000283
<input checked="" type="checkbox"/>	Chain	20 – 177	158	Urokinase-type plasminogen activator long chain A <small>By similarity</small>		PRO_00000283
<input checked="" type="checkbox"/>	Chain	156 – 177	22	Urokinase-type plasminogen activator short chain A <small>By similarity</small>		PRO_00000283
<input checked="" type="checkbox"/>	Chain	179 – 432	254	Urokinase-type plasminogen activator chain B <small>By similarity</small>		PRO_00000283

**Regions**

<input checked="" type="checkbox"/>	Domain	27 – 63	37	EGF-like		
<input checked="" type="checkbox"/>	Domain	70 – 151	82	Kringle		
<input checked="" type="checkbox"/>	Domain	179 – 425	247	Peptidase S1		
<input checked="" type="checkbox"/>	Region	34 – 57	24	Binds urokinase plasminogen activator surface receptor		
<input checked="" type="checkbox"/>	Region	152 – 178	27	Connecting peptide		

**Sites**

<input type="checkbox"/>	Active site	225	1	Charge relay system		
<input type="checkbox"/>	Active site	276	1	Charge relay system		
<input type="checkbox"/>	Active site	377	1	Charge relay system		

**Amino acid modifications**

<input type="checkbox"/>	Disulfide bond	31 ↔ 39		<small>By similarity</small>		
<input type="checkbox"/>	Disulfide bond	33 ↔ 51		<small>By similarity</small>		



<input type="checkbox"/>	Disulfide bond	53 ↔ 62		By similarity	■	
<input type="checkbox"/>	Disulfide bond	70 ↔ 151		By similarity	■ ■	
<input type="checkbox"/>	Disulfide bond	91 ↔ 133		By similarity	■ ■	
<input type="checkbox"/>	Disulfide bond	122 ↔ 146		By similarity	■ ■	
<input type="checkbox"/>	Disulfide bond	168 ↔ 300		Interchain (between A and B chains) By similarity	■ ■	
<input type="checkbox"/>	Disulfide bond	210 ↔ 226		By similarity	■ ■	
<input type="checkbox"/>	Disulfide bond	218 ↔ 289		By similarity	■ ■	
<input type="checkbox"/>	Disulfide bond	314 ↔ 383		By similarity	■ ■	
<input type="checkbox"/>	Disulfide bond	346 ↔ 362		By similarity	■ ■	
<input type="checkbox"/>	Disulfide bond	373 ↔ 401		By similarity	■ ■	

Experimental info

<input type="checkbox"/>	Sequence conflict	16	1	N → H in CAA46601. Ref.2	■	
<input type="checkbox"/>	Sequence conflict	24	1	E → G in CAA46601. Ref.2	■	
<input type="checkbox"/>	Sequence conflict	332	1	D → N in CAA46601. Ref.2	■	

Sequences

Sequence	Length	Mass (Da)
<input type="checkbox"/> P29598 [UniParc]. Last modified April 1, 1993. Version 1. Checksum: 4EB1B96C716244C8	FASTA 432	47,957
<div><div><div>102030405060</div><div>MRVWLASLFL CALVANSEGG SELEASDESN CGCQNGGVCV SYKYFSSIRR CSCPCKFKGE</div></div><div><div>708090100110120</div><div>HCEIDTSKTC YHGNGQSYRG KANTDTKGRP CLAWNPAVL QQTYNHRSD ALSLGLGKHN</div></div><div><div>130140150160170180</div><div>YCRNPDNQRR PWCYVQIGLK QFVQECMVQD CSLSKKPSST VDQQGFQCGQ KALRPRFKIV</div></div><div><div>190200210220230240</div><div>GGEFTVVENQ PWFAAIYLKN KGGSPPSFKC GGSLISPCWV ASATHCFVNQ PKKEEYVVYL</div></div><div><div>250260270280290300</div><div>GQSKRNSYNP GEMKFEVEQL ILHEDFSDET LAFHNDIALL KIRTSTGQCA QPSRTIQTIC</div></div><div><div>310320330340350360</div><div>LPPRFGDAPF GSDCEITGFG QESATDYFYP KDLKMSVVKI ISHEQCKQPH YYGSEINYKM</div></div><div><div>370380390400410420</div><div>LCAADPEWKT DSCSGDSGGP LICNIDGRPT LSGIVSWGSG CAEKNKPGVY TRVSYFLNWI</div></div></div>		

430  
QSHIGEENGL AF

« Hide

## References

- |     |   |
|-----|---|
| [1] | <p><b>"Transcriptional and posttranscriptional activation of urokinase plasminogen activator gene expression in metastatic tumor cells."</b><br/> Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A., Kefford R.F.<br/> Cancer Res. 52:2489-2496(1992) [PubMed: 1568219] [Abstract]<br/> <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].<br/> <u>Strain:</u> Fischer 344.</p>                                       |
| [2] | <p>Rabbani S.A.<br/> Submitted (APR-1992) to the EMBL/GenBank/DDBJ databases<br/> <u>Cited for:</u> NUCLEOTIDE SEQUENCE [MRNA].<br/> <u>Tissue:</u> Kidney.</p>   |
| [3] | <p><b>"The receptor for the plasminogen activator of urokinase type is up-regulated in transformed rat thyroid cells."</b><br/> Ragno P., Cassano S., Degen J., Kessler C., Blasi F., Rossi G.<br/> FEBS Lett. 306:193-198(1992) [PubMed: 1321734] [Abstract]<br/> <u>Cited for:</u> NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 31-62, IDENTIFICATION OF UROKINASE PLASMINOGEN ACTIVATOR SURFACE RECEPTOR BINDING DOMAIN.</p> |
| +   | Additional computationally mapped references.   |

## Cross-references

### Sequence databases

<input checked="" type="radio"/> EMBL <input checked="" type="radio"/> GenBank <input checked="" type="radio"/> DDBJ	X63434 mRNA. Translation: CAA45028.1. X65651 mRNA. Translation: CAA46601.1. X66907 Genomic DNA. Translation: CAA47356.1.
IPI	IPI00205838.
PIR	S18932. S24604.
UniGene	Rn.6064.

### 3D structure databases

ProteinModelPortal	P29598.
SMR	P29598. Positions 29-155, 166-427.
ModBase	Search...

### Protein-protein interaction databases

STRING	P29598.
--------	---------

### Protein family/group databases

MEROPS	S01.231.
--------	----------

<b>PTM databases</b>	
PhosphoSite	P29598.
<b>Genome annotation databases</b>	
Ensembl	ENSRNOT00000014273; ENSRNOP00000014273; ENSRNOG00000010516.
<b>Organism-specific databases</b>	
RGD	3343. Plau.
<b>Phylogenomic databases</b>	
eggNOG	roNOG07876.
HOVERGEN	HBG008633.
InParanoid	P29598.
OrthoDB	EOG4229K6.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.21.73. 248.
<b>Gene expression databases</b>	
ArrayExpress	P29598.
Genevestigator	P29598.
GermOnline	ENSRNOG00000010516. Rattus norvegicus.
<b>Family and domain databases</b>	
InterPro	IPR006210. EGF-like. IPR013032. EGF-like_reg_CS. IPR000742. EGF_3. IPR000001. Kringle. IPR013806. Kringle-like. IPR018056. Kringle_CS. IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Gene3D	G3DSA:2.40.20.10. Kringle. 1 hit.
Pfam	PF00051. Kringle. 1 hit. PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN. PR00018. KRINGLE.
SMART	SM00181. EGF. 1 hit. SM00130. KR. 1 hit. SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF57440. Kringle-like. 1 hit. SSF50494. Pept_Ser_Cys. 1 hit.

PROSITE	PS00022. EGF_1. 1 hit. PS01186. EGF_2. False negative. PS50026. EGF_3. 1 hit. PS00021. KRINGLE_1. 1 hit. PS50070. KRINGLE_2. 1 hit. PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. False negative. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Entry information</b>	
Entry name	UROK_RAT
Accession	Primary (citable) accession number: <b>P29598</b> Secondary accession number(s): Q6LBK5
Entry history	Integrated into      April 1, 1993 UniProtKB/Swiss- Prot: Last sequence      April 1, 1993 update: Last modified:      April 5, 2011 This is version 103 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Relevant documents</b>	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License &amp; Disclaimer | Contact





**Q05589 (UROK\_BOVIN) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 97.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Urokinase-type plasminogen activator</b> Short name=U-plasminogen activator Short name=uPA EC=3.4.21.73 <i>Cleaved into the following 3 chains:</i> 1. <b>Urokinase-type plasminogen activator long chain</b> A 2. <b>Urokinase-type plasminogen activator short chain</b> A 3. <b>Urokinase-type plasminogen activator chain B</b>
Gene names	Name:PLAU
Organism	<b>Bos taurus (Bovine)</b>
Taxonomic identifier	9913 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Laurasiatheria › Cetartiodactyla › Ruminantia › Pecora › Bovidae › Bovinae › Bos

**Protein attributes**

Sequence length	433 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at transcript level.

**General annotation (Comments)**

Function	Specifically cleave the zymogen plasminogen to form the active enzyme plasmin.
Catalytic activity	Specific cleavage of Arg- -Val bond in plasminogen to form plasmin.
Subunit structure	Found in high and low molecular mass forms. Each consists of two chains, A and B. The high molecular mass form contains a long chain A which is cleaved to yield a short chain A. Binds LRP1B; binding is followed by internalization and degradation. Interacts with MRC2. Interacts with PLAUR <span>By similarity</span> .
Subcellular location	Secreted.
Induction	By retinoic acid.
Sequence similarities	Belongs to the peptidase S1 family. Contains 1 EGF-like domain.

**P16227 (UROK\_PAPCY)** ★ Reviewed, UniProtKB/Swiss-Prot

Last modified January 11, 2011. Version 92.

**Names and origin**

Protein names	<p><i>Recommended name:</i>  <b>Urokinase-type plasminogen activator</b>  Short name=U-plasminogen activator  Short name=uPA  EC=3.4.21.73</p> <p><u>Cleaved into the following 3 chains:</u></p> <ol style="list-style-type: none"> <li>1. <b><i>Urokinase-type plasminogen activator long chain A</i></b></li> <li>2. <b><i>Urokinase-type plasminogen activator short chain A</i></b></li> <li>3. <b><i>Urokinase-type plasminogen activator chain B</i></b></li> </ol>
Gene names	Name:PLAU
Organism	Papio cynocephalus (Yellow baboon)
Taxonomic identifier	9556 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Cercopithecidae › Cercopithecinae › Papio

**Protein attributes**

Sequence length	433 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at transcript level.

**General annotation (Comments)**

Function	Specifically cleave the zymogen plasminogen to form the active enzyme plasmin.
Catalytic activity	Specific cleavage of Arg- -Val bond in plasminogen to form plasmin.
Subunit structure	Found in high and low molecular mass forms. Each consists of two chains, A and B. The high molecular mass form contains a long chain A which is cleaved to yield a short chain A. Binds LRP1B; binding is followed by internalization and degradation. Interacts with MRC2. Interacts with PLAUR <span>By similarity</span> .
Subcellular location	Secreted.
Post-translational modification	Phosphorylation of Ser-157 and Ser-325 abolishes proadhesive ability but does not interfere with receptor binding <span>By similarity</span> .
Sequence similarities	Belongs to the peptidase S1 family.

	Contains 1 EGF-like domain. Contains 1 kringle domain. Contains 1 peptidase S1 domain.
--	--

Ontologies

Keywords




Biological process	Plasminogen activation
Cellular component	Secreted
Domain	EGF-like domain Kringle Signal
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Glycoprotein Phosphoprotein Zymogen



Gene Ontology (GO)

Biological process	proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	extracellular region Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	serine-type endopeptidase activity Inferred from electronic annotation. Source: UniProtKB-KW






Complete GO annotation...

Sequence annotation (Features)




Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
Molecule processing					
 Signal peptide	1 – 20	20	Potential		
 Chain	21 – 433	413	Urokinase-type plasminogen activator		PRO_0000028
 Chain	21 – 176	156	Urokinase-type plasminogen activator long chain A By similarity		PRO_0000028

	Chain	155 – 176	22	Urokinase-type plasminogen activator short chain A <small>By similarity</small>		PRO_0000028
	Chain	178 – 433	256	Urokinase-type plasminogen activator chain B <small>By similarity</small>		PRO_0000028









**Regions**

	Domain	26 – 62	37	EGF-like		
	Domain	69 – 150	82	Kringle		
	Domain	178 – 426	249	Peptidase S1		
	Region	33 – 56	24	Binds urokinase plasminogen activator surface receptor <small>By similarity</small>		
	Region	151 – 177	27	Connecting peptide		

**Sites**

	Active site	223	1	Charge relay system		
	Active site	274	1	Charge relay system		
	Active site	378	1	Charge relay system		

**Amino acid modifications**

	Modified residue	157	1	Phosphoserine <small>By similarity</small>		
	Modified residue	325	1	Phosphoserine <small>By similarity</small>		
	Glycosylation	324	1	N-linked (GlcNAc...) <small>Potential</small>		
	Disulfide bond	30 ↔ 38		<small>By similarity</small>		
	Disulfide bond	32 ↔ 50		<small>By similarity</small>		
	Disulfide bond	52 ↔ 61		<small>By similarity</small>		
	Disulfide bond	69 ↔ 150		<small>By similarity</small>		
	Disulfide bond	90 ↔ 132		<small>By similarity</small>		



<input type="checkbox"/>	Disulfide bond	121 ↔ 145	<input type="checkbox"/> By similarity		
<input type="checkbox"/>	Disulfide bond	167 ↔ 298	Interchain (between A and B chains) <input type="checkbox"/> By similarity		
<input type="checkbox"/>	Disulfide bond	208 ↔ 224	<input type="checkbox"/> By similarity		
<input type="checkbox"/>	Disulfide bond	216 ↔ 287	<input type="checkbox"/> By similarity		
<input type="checkbox"/>	Disulfide bond	315 ↔ 384	<input type="checkbox"/> By similarity		
<input type="checkbox"/>	Disulfide bond	347 ↔ 363	<input type="checkbox"/> By similarity		
<input type="checkbox"/>	Disulfide bond	374 ↔ 402	<input type="checkbox"/> By similarity		

Sequences

Sequence	Length	Mass (Da)
<input type="checkbox"/> P16227 [UniParc]. Last modified April 1, 1990. Version 1. Checksum: 816D22DFEDDC8792	FASTA 433	48,595
<div><div><div>102030405060</div><div>MRALLAHL<sup>10</sup>LL<sup>20</sup>CVLVVSASK<sup>30</sup>G<sup>40</sup>SRELQVPSDC<sup>50</sup>GCLNGGTCMS<sup>60</sup>NKYFSSIHWC<sup>70</sup>NCPKKFGGQH<sup>80</sup></div></div><div><div>708090100110120</div><div>CEIDKSKTCY<sup>70</sup>EGNGHFYRGK<sup>80</sup>ASTDTMGRSC<sup>90</sup>LAWNSATVLQ<sup>100</sup>QTYHAHRSDA<sup>110</sup>LQLGLGKHNY<sup>120</sup></div></div><div><div>130140150160170180</div><div>CRNPDNRRRP<sup>130</sup>WCYVQVGLKQ<sup>140</sup>RVQECMVHNC<sup>150</sup>ADGKKPSSPP<sup>160</sup>EELQFQCGQR<sup>170</sup>TLRPRFKIVG<sup>180</sup></div></div><div><div>190200210220230240</div><div>GEFTTIENQP<sup>190</sup>WFAAIYRRHR<sup>200</sup>GGSVTYVCGG<sup>210</sup>SLISPCWVVS<sup>220</sup>ATHCFINYPK<sup>230</sup>KEDYIVYLGR<sup>240</sup></div></div><div><div>250260270280290300</div><div>SRLNSNTQGE<sup>250</sup>MKFEVENLIL<sup>260</sup>HEDYSADTLA<sup>270</sup>HHNDIALSKI<sup>280</sup>RSKEGRCAQP<sup>290</sup>SRTIQTICLP<sup>300</sup></div></div><div><div>310320330340350360</div><div>SMYNDPNDPP<sup>310</sup>FGTSCEITGF<sup>320</sup>GKENSTDYLY<sup>330</sup>PEQLKMTVVK<sup>340</sup>LVSHQKCQQP<sup>350</sup>HYYGSEVTTK<sup>360</sup></div></div><div><div>370380390400410420</div><div>MLCAADPQWE<sup>370</sup>TDSCQGDSSG<sup>380</sup>PLVCSIQGHM<sup>390</sup>TLTGIVSWGR<sup>400</sup>GCALKDKPGV<sup>410</sup>YTRVSRFLPW<sup>420</sup></div></div><div><div>430</div><div>IHSHTREQNG<sup>430</sup>LAL</div></div></div>		

« Hide

References

[1] "Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen activator."  
Au Y.P.T., Wang T.W., Clowes A.W.  
Nucleic Acids Res. 18:3411-3411(1990) [PubMed: 2113276] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [MRNA].  
Tissue: Thoracic aorta.

**Cross-references****Sequence databases**

<ul style="list-style-type: none"> <li>⊙ EMBL</li> <li>⊙ GenBank</li> <li>⊙ DDBJ</li> </ul>	X51935 mRNA. Translation: CAA36200.1.
---	---------------------------------------

PIR	UKBAY. S14687.
-----	----------------

**3D structure databases**

ProteinModelPortal	P16227.
--------------------	---------

SMR	P16227. Positions 25-154, 165-428.
-----	------------------------------------

ModBase	Search...
---------	-----------

**Protein family/group databases**

MEROPS	S01.231.
--------	----------

**Phylogenomic databases**

HOVERGEN	HBG008633.
----------	------------

**Enzyme and pathway databases**

BRENDA	3.4.21.73. 74251.
--------	-------------------

**Family and domain databases**

InterPro	IPR006210. EGF-like. IPR013032. EGF-like_reg_CS. IPR000742. EGF_3. IPR000001. Kringle. IPR013806. Kringle-like. IPR018056. Kringle_CS. IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
----------	--

Gene3D	G3DSA:2.40.20.10. Kringle. 1 hit.
--------	-----------------------------------

Pfam	PF00051. Kringle. 1 hit. PF00089. Trypsin. 1 hit. [Graphical view]
------	--

PRINTS	PR00722. CHYMOTRYPSIN. PR00018. KRINGLE.
--------	---

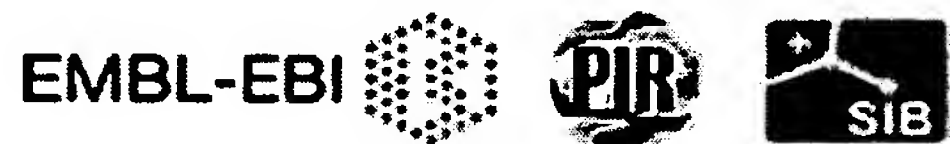
SMART	SM00181. EGF. 1 hit. SM00130. KR. 1 hit. SM00020. Tryp_SPc. 1 hit. [Graphical view]
-------	--

SUPFAM	SSF57440. Kringle-like. 1 hit. SSF50494. Pept_Ser_Cys. 1 hit.
--------	--

PROSITE	PS00022. EGF_1. 1 hit. PS01186. EGF_2. False negative. PS50026. EGF_3. 1 hit.
---------	---

ProtoNet	PS00021. KRINGLE_1. 1 hit. PS50070. KRINGLE_2. 1 hit. PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
	Search...
<b>Entry information</b>	
Entry name	UROK_PAPCY
Accession	Primary (citable) accession number: <b>P16227</b>
Entry history	Integrated into April 1, 1990 UniProtKB/Swiss-Prot: Last sequence April 1, 1990 update: Last modified: January 11, 2011 This is version 92 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Relevant documents</b>	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	

© 2002–2011 UniProt Consortium | License &amp; Disclaimer | Contact



**P04185 (UROK\_PIG) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 94.

**Names and origin**

Protein names	<b>Recommended name:</b> <b>Urokinase-type plasminogen activator</b> Short name=U-plasminogen activator Short name=uPA EC=3.4.21.73 <b>Cleaved into the following 3 chains:</b> 1. <b>Urokinase-type plasminogen activator long chain A</b> 2. <b>Urokinase-type plasminogen activator short chain A</b> 3. <b>Urokinase-type plasminogen activator chain B</b>
Gene names	Name:PLAU
Organism	<b>Sus scrofa (Pig)</b>
Taxonomic identifier	9823 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Laurasiatheria › Cetartiodactyla › Suina › Suidae › Sus

**Protein attributes**

Sequence length	442 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at transcript level.

**General annotation (Comments)**

Function	Specifically cleave the zymogen plasminogen to form the active enzyme plasmin.
Catalytic activity	Specific cleavage of Arg- -Val bond in plasminogen to form plasmin.
Subunit structure	Found in high and low molecular mass forms. Each consists of two chains, A and B. The high molecular mass form contains a long chain A which is cleaved to yield a short chain A. Binds LRP1B; binding is followed by internalization and degradation. Interacts with MRC2. Interacts with PLAUR <a href="#">By similarity</a> .
Subcellular location	Secreted <a href="#">By similarity</a> .
Sequence similarities	Belongs to the peptidase S1 family. Contains 1 EGF-like domain. Contains 1 kringle domain.



Contains 1 peptidase S1 domain.

Ontologies

Keywords





Biological process	Plasminogen activation
Cellular component	Secreted
Domain	EGF-like domain Kringle Signal
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Glycoprotein Zymogen

Gene Ontology (GO)

Biological process	proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	extracellular region Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	serine-type endopeptidase activity Inferred from electronic annotation. Source: UniProtKB-KW

Complete GO annotation...

Sequence annotation (Features)

	Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
Molecule processing						
	Signal peptide	1 – 20	20	<div>By similarity</div>		
	Chain	21 – 442	422	Urokinase-type plasminogen activator		PRO_00000283
	Chain	21 – 188	168	Urokinase-type plasminogen activator long chain A <div>By similarity</div>		PRO_00000283
	Chain	167 – 188	22	Urokinase-type plasminogen activator		PRO_00002859

<input checked="" type="checkbox"/>	Chain	190 – 442	253	short chain A By similarity Urokinase-type plasminogen activator chain B By similarity		PRO_00000283
<b>Regions</b>						
<input checked="" type="checkbox"/>	Domain	29 – 65	37	EGF-like		
<input checked="" type="checkbox"/>	Domain	72 – 153	82	Kringle		
<input checked="" type="checkbox"/>	Domain	190 – 435	246	Peptidase S1		
<input checked="" type="checkbox"/>	Region	36 – 59	24	Binds urokinase plasminogen activator surface receptor By similarity		
<input checked="" type="checkbox"/>	Region	154 – 189	36	Connecting peptide		
<b>Sites</b>						
<input type="checkbox"/>	Active site	235	1	Charge relay system		
<input type="checkbox"/>	Active site	286	1	Charge relay system		
<input type="checkbox"/>	Active site	387	1	Charge relay system		
<b>Amino acid modifications</b>						
<input type="checkbox"/>	Glycosylation	152	1	N-linked (GlcNAc...) Potential		
<input type="checkbox"/>	Disulfide bond	33 ↔ 41		By similarity		
<input type="checkbox"/>	Disulfide bond	35 ↔ 53		By similarity		
<input type="checkbox"/>	Disulfide bond	55 ↔ 64		By similarity		
<input type="checkbox"/>	Disulfide bond	72 ↔ 153		By similarity		
<input type="checkbox"/>	Disulfide bond	93 ↔ 135		By similarity		
<input type="checkbox"/>	Disulfide bond	124 ↔ 148		By similarity		
<input type="checkbox"/>	Disulfide bond	179 ↔ 310		Interchain (between A and B)		

			chains)	
			By similarity	
<input type="checkbox"/>	Disulfide bond	220 ↔ 236	By similarity	
<input type="checkbox"/>	Disulfide bond	228 ↔ 299	By similarity	
<input type="checkbox"/>	Disulfide bond	324 ↔ 393	By similarity	
<input type="checkbox"/>	Disulfide bond	356 ↔ 372	By similarity	
<input type="checkbox"/>	Disulfide bond	383 ↔ 411	By similarity	

Experimental info

<input type="checkbox"/>	Sequence conflict	241	1	Q → H in CAA25806. Ref.1	
<input type="checkbox"/>	Sequence conflict	242	1	Q → H in CAA26511. Ref.1	
<input type="checkbox"/>	Sequence conflict	288	1	A → GS in CAA25806. Ref.1	

Sequences

Sequence	Length	Mass (Da)
<input type="checkbox"/> P04185 [UniParc]. Last modified August 13, 1987. Version 1. Checksum: EE32FCEF501321EE	FASTA	442 49,117

10	20	30	40	50	60
MRVLRACLSL	CVLVVS	DSKGS	SHELHQESGA	SNCGCLNGGK	CVSYKYFSNI
QRCSCP	KKFQ				
70	80	90	100	110	120
GEHCEIDTSQ	TCFEGNGHSY	RGKANTNTGG	RPCLPWNSAT	VLLNTYHAHR	PDALQLGLGK
130	140	150	160	170	180
HNYCRNPDNQ	RRPWCYVQVG	LKQLVQECMV	PNCSGGESHR	PAYDGKNPFS	TPEKVEFQCG
190	200	210	220	230	240
QKALRPRFKI	VGGKSTTIEN	QPWFAAIYRR	HRGGSVTYVC	GGSLISPCWV	VSATHCFINY
250	260	270	280	290	300
QQKEDIIVYL	GRQTLHSSTH	GEMKFEVEKL	ILHEDYSADS	LAHNDIALL	KIRTDKGQCA
310	320	330	340	350	360
QPSRSIQTIC	LPPVNGDAHF	GASCEIVGFG	KEDPSDYLYP	EQLKMTTVVKL	VSHRECQQPH
370	380	390	400	410	420
YYGSEVTTKM	LCAADPQWKT	DSCQGDGGP	LVCSTQGRLT	LTGIVSWGRE	CAMKDKPGVY
430	440				
TRVSRFLTWI	HTHVGGENGL	AH			

« Hide

References

[1]	"cDNA and gene nucleotide sequence of porcine plasminogen activator." Nagamine Y., Pearson D., Altus M.S., Reich E.
-----	--

- Nucleic Acids Res. 12:9525-9541(1984) [PubMed: 6096832] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [GENOMIC DNA / MRNA].  
Tissue: Kidney.
- [2] Nagamine Y.  
 Submitted (DEC-1986) to the PIR data bank  
Cited for: SEQUENCE REVISION TO 241.

## Cross-references

### Sequence databases

<ul style="list-style-type: none"> <li>⊙ EMBL</li> <li>⊙ GenBank</li> <li>⊙ DDBJ</li> </ul>	X01648 Genomic DNA. Translation: CAA25806.1. X02724 mRNA. Translation: CAA26511.1.
---	---

PIR	UKPG. A00932.
RefSeq	NP_999110.1. NM_213945.1.
UniGene	Ssc.11194.

### 3D structure databases

ProteinModelPortal	P04185.
SMR	P04185. Positions 31-157, 177-437.
ModBase	Search...

### Protein family/group databases

MEROPS	S01.231.
--------	----------

### Genome annotation databases

Ensembl	ENSSSCT00000011287; ENSSSCP00000010995; ENSSSCG00000010312.
GeneID	396985.
KEGG	ssc:396985.

### Organism-specific databases

CTD	396985.
-----	---------

### Phylogenomic databases

GeneTree	ENSGT00570000078945.
HOVERGEN	HBG008633.
OrthoDB	EOG4229K6.

### Enzyme and pathway databases

BRENDA	3.4.21.73. 249.
--------	-----------------



**Family and domain databases**

InterPro	IPR013032. EGF-like_reg_CS. IPR000742. EGF_3. IPR000001. Kringle. IPR013806. Kringle-like. IPR018056. Kringle_CS. IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Gene3D	G3DSA:2.40.20.10. Kringle. 1 hit.
Pfam	PF00051. Kringle. 1 hit. PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN. PR00018. KRINGLE.
SMART	SM00130. KR. 1 hit. SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF57440. Kringle-like. 1 hit. SSF50494. Pept_Ser_Cys. 1 hit.
PROSITE	PS00022. EGF_1. 1 hit. PS01186. EGF_2. False negative. PS50026. EGF_3. 1 hit. PS00021. KRINGLE_1. 1 hit. PS50070. KRINGLE_2. 1 hit. PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...

**Entry information**

Entry name	UROK_PIG
Accession	Primary (citable) accession number: <b>P04185</b>
Entry history	Integrated into March 20, 1987 UniProtKB/Swiss-Prot: Last sequence August 13, 1987 update: Last modified: April 5, 2011 This is version 94 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program

**Relevant documents**

Peptidase families  
Classification of peptidase families and list of entries

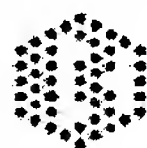
---

SIMILARITY comments

Index of protein domains and families

© 2002–2011 UniProt Consortium | License & Disclaimer | Contact

EMBL-EBI



**P15120 (UROK\_CHICK) ★ Reviewed, UniProtKB/Swiss-Prot**

Last modified April 5, 2011. Version 101.

**Names and origin**

Protein names	<i>Recommended name:</i> <b>Urokinase-type plasminogen activator</b> Short name=U-plasminogen activator Short name=uPA EC=3.4.21.73 <i>Cleaved into the following 2 chains:</i> 1. <b>Urokinase-type plasminogen activator chain A</b> 2. <b>Urokinase-type plasminogen activator chain B</b>
Gene names	Name: <b>PLAU</b>
Organism	<b>Gallus gallus (Chicken)</b>
Taxonomic identifier	9031 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Archosauria › Dinosauria › Saurischia › Theropoda › Coelurosauria › Aves › Neognathae › Galliformes › Phasianidae › Phasianinae › Gallus

**Protein attributes**

Sequence length	434 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at transcript level.






**General annotation (Comments)**

Function	Specifically cleave the zymogen plasminogen to form the active enzyme plasmin.
Catalytic activity	Specific cleavage of Arg-I-Val bond in plasminogen to form plasmin.
Subcellular location	Secreted <span>By similarity</span> .
Sequence similarities	Belongs to the peptidase S1 family. Contains 1 EGF-like domain. Contains 1 kringle domain. Contains 1 peptidase S1 domain.





**Ontologies****Keywords**

Biological process	Plasminogen activation
Cellular component	Secreted
Domain	EGF-like domain Kringle Signal
Molecular function	Hydrolase Protease Serine protease
PTM	Disulfide bond Glycoprotein Zymogen
<b>Gene Ontology (GO)</b>	
Biological process	proteolysis Inferred from electronic annotation. Source: InterPro
Cellular component	extracellular region Inferred from electronic annotation. Source: UniProtKB-SubCell
Molecular function	serine-type endopeptidase activity Inferred from electronic annotation. Source: UniProtKB-KW
Complete GO annotation...	

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
<b>Molecule processing</b>					
 Signal peptide	1 – 20	20	Potential		
 Chain	21 – 434	414	Urokinase-type plasminogen activator		PRO_0000028337
 Chain	21 – 171	151	Urokinase-type plasminogen activator chain A By similarity		PRO_0000028338
 Chain	173 – 434	262	Urokinase-type plasminogen activator chain B By similarity		PRO_0000028339
<b>Regions</b>					
 Domain	36 – 72	37	EGF-like		



	Domain	79 – 158	80	Kringle		
	Domain	173 – 421	249	Peptidase S1		
	Region	159 – 172	14	Connecting peptide		
Sites						
<input type="checkbox"/>	Active site	217	1	Charge relay system <div>By similarity</div>		
<input type="checkbox"/>	Active site	272	1	Charge relay system <div>By similarity</div>		
<input type="checkbox"/>	Active site	373	1	Charge relay system <div>By similarity</div>		
Amino acid modifications						
<input type="checkbox"/>	Glycosylation	228	1	N-linked (GlcNAc...) <div>Potential</div>		
<input type="checkbox"/>	Disulfide bond	40 ↔ 48		<div>By similarity</div>		
<input type="checkbox"/>	Disulfide bond	42 ↔ 60		<div>By similarity</div>		
<input type="checkbox"/>	Disulfide bond	62 ↔ 71		<div>By similarity</div>		
<input type="checkbox"/>	Disulfide bond	79 ↔ 158		<div>By similarity</div>		
<input type="checkbox"/>	Disulfide bond	96 ↔ 139		<div>By similarity</div>		
<input type="checkbox"/>	Disulfide bond	128 ↔ 152		<div>By similarity</div>		
<input type="checkbox"/>	Disulfide bond	162 ↔ 296		Interchain (between A and B chains) <div>By similarity</div>		
<input type="checkbox"/>	Disulfide bond	202 ↔ 218		<div>By similarity</div>		
<input type="checkbox"/>	Disulfide bond	210 ↔ 285		<div>By similarity</div>		
<input type="checkbox"/>	Disulfide bond	310 ↔ 379		<div>By similarity</div>		
<input type="checkbox"/>	Disulfide bond	342 ↔ 358		<div>By similarity</div>		
<input type="checkbox"/>	Disulfide bond	369 ↔ 397		<div>By similarity</div>		
Sequences						
<div><div></div><div>Sequence</div><div>Length</div><div>Mass (Da)</div></div>						
	P15120 [UniParc]. Last modified April 1, 1990. Version 1. Checksum: BD881048DD666A55	FASTA	434	49,400		

10	20	30	40	50	60
MKLIIFLTVT	LCTLVTGLDS	VYIRQYYKLS	HKHRPQHREC	QCLNGGTCIT	YRFFSQIKRC
70	80	90	100	110	120
LCPEGYGGLH	CEIDTNSICY	SGNGEDYRGM	AEDPGCLYWD	HPSVIRWGDY	HADLKNALQL
130	140	150	160	170	180
GLGKHNYCRN	PNGRSRPWCY	TKRRYSIQET	PCSTIEKCER	TCGQRSFSKY	FKIVGGSQAE
190	200	210	220	230	240
VETQPWIAGI	FQNIMGTDQF	LCGGSLIDPC	WVLTAAHCFY	NPTKKQPNKS	VYKVFLGKSI
250	260	270	280	290	300
LNTNDEHEQV	FMVDEIISHP	DFTDHTGGND	NDIALIRIRT	ASGQCAVESN	YVRTVCLPEK
310	320	330	340	350	360
NLNLYDNTWC	EIAGYGKQNS	YDIYYAQRML	SATVNLISQD	DCKNKYDST	RVTDNMVCAG
370	380	390	400	410	420
DPLWETDACK	GDSGGPMVCE	HNGRMTLYGI	VSWGDCAKK	NKPGVYTRVT	RYLNWIDSNM
430					
NAVFTKSRSF	REPK				

« Hide

References

[1] "The chicken urokinase-type plasminogen activator gene."  
Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.  
J. Biol. Chem. 265:1339-1344(1990) [PubMed: 2295632] [Abstract]  
Cited for: NUCLEOTIDE SEQUENCE [GENOMIC DNA / MRNA].

Cross-references

Sequence databases

Ⓢ EMBL	J05187 mRNA. Translation: AAA49131.1.
Ⓢ GenBank	J05188 Genomic DNA. Translation: AAA49130.1.
Ⓢ DDBJ	

IPI	IPI00584377.
PIR	A35005.
RefSeq	NP_990774.2. NM_205443.2.
UniGene	Gga.817.

3D structure databases

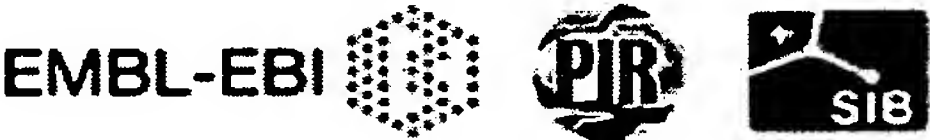
ProteinModelPortal	P15120.
SMR	P15120. Positions 60-422.
ModBase	Search...

Protein-protein interaction databases

STRING	P15120.
--------	---------

<b>Protein family/group databases</b>	
MEROPS	S01.232.
<b>Genome annotation databases</b>	
Ensembl	ENSGALT00000008157; ENSGALP00000008143; ENSGALG00000005086.
GeneID	396424.
KEGG	gga:396424.
<b>Organism-specific databases</b>	
CTD	396424.
<b>Phylogenomic databases</b>	
eggNOG	veNOG07790.
GeneTree	ENSGT00570000078945.
HOGENOM	HBG755338.
HOVERGEN	HBG008633.
InParanoid	P15120.
OrthoDB	EOG4229K6.
<b>Enzyme and pathway databases</b>	
BRENDA	3.4.21.73. 4.
<b>Family and domain databases</b>	
InterPro	IPR006210. EGF-like. IPR013032. EGF-like_reg_CS. IPR000742. EGF_3. IPR000001. Kringle. IPR013806. Kringle-like. IPR018056. Kringle_CS. IPR009003. Pept_cys/ser_Trypsin-like. IPR018114. Peptidase_S1/S6_AS. IPR001254. Peptidase_S1_S6. IPR001314. Peptidase_S1A. [Graphical view]
Gene3D	G3DSA:2.40.20.10. Kringle. 1 hit.
Pfam	PF00051. Kringle. 1 hit. PF00089. Trypsin. 1 hit. [Graphical view]
PRINTS	PR00722. CHYMOTRYPSIN. PR00018. KRINGLE.
SMART	SM00181. EGF. 1 hit. SM00130. KR. 1 hit. SM00020. Tryp_SPc. 1 hit. [Graphical view]
SUPFAM	SSF57440. Kringle-like. 1 hit. SSF50494. Pept_Ser_Cys. 1 hit.

PROSITE	PS00022. EGF_1. 1 hit. PS01186. EGF_2. 1 hit. PS50026. EGF_3. 1 hit. PS00021. KRINGLE_1. 1 hit. PS50070. KRINGLE_2. 1 hit. PS50240. TRYPSIN_DOM. 1 hit. PS00134. TRYPSIN_HIS. 1 hit. PS00135. TRYPSIN_SER. 1 hit. [Graphical view]
ProtoNet	Search...
<b>Entry information</b>	
Entry name	UROK_CHICK
Accession	Primary (citable) accession number: <b>P15120</b>
Entry history	Integrated into April 1, 1990 UniProtKB/Swiss-Prot: Last sequence April 1, 1990 update: Last modified: April 5, 2011 This is version 101 of the entry and version 1 of the sequence. [Complete history]
Entry status	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Chordata Protein Annotation Program
<b>Relevant documents</b>	
Peptidase families Classification of peptidase families and list of entries	
SIMILARITY comments Index of protein domains and families	





# Comparative Modeling Methods: Application to the Family of the Mammalian Serine Proteases

Jonathan Greer

*Computer Assisted Molecular Design Group, Pharmaceutical Products Division, Abbott Laboratories, Abbott Park, Illinois 60064*

**ABSTRACT** Comparative modeling methods are described that can be used to construct a three-dimensional model structure of a new protein from knowledge of its sequence and of the experimental structures and sequences of other members of its homology family. The methods are illustrated with the mammalian serine protease family, for which seven experimental structures have been reported in the literature, and the sequences for over 35 different protein members of the family are available. The strategy for modeling these proteins is presented, and criteria are developed for determining and assigning the reliability of the modeled structure. Criteria are described that are specially designed to help detect cases in which it is likely that the local structure diverges significantly from the usual conformation of the family.

**Key words:** protein structure, computer modeling, structure prediction, sequence homology, structure homology

## INTRODUCTION

It has been apparent for close to two decades that proteins from very different sources and sometimes with rather diverse functions can have homologous sequences and consequently very similar three-dimensional structures.<sup>1</sup> This fact has been the basis for the development of comparative modeling methods,<sup>2-7</sup> which permit extrapolation from the experimentally determined structure for one or more members of a homologous family to a new member of this family whose sequence has been determined but whose structure is as yet unknown. A number of factors combine to increase greatly the application of comparative modeling techniques today. The large number of protein structures<sup>8</sup> and the exploding number of protein sequences<sup>9</sup> that are being reported in the literature and that are readily available in computerized databases provide the basic structural and sequence data needed to apply the method. The proteins in which we are interested are often available in only small quantities, too little for structural studies unless the gene or mRNA is cloned or synthesized and expressed. Even when this latter effort is deemed worthwhile and is initiated,

the comparative modeling studies can be performed in the meantime, providing an approximate view of the structure of the molecule until sufficient protein can be obtained and the experimental structure can be determined. Such a model structure can be very useful in the interim to help plan and interpret biochemical<sup>10</sup> and mutagenesis<sup>11</sup> experiments, to probe functional properties,<sup>12</sup> and improve our understanding of ligand or substrate binding.<sup>13-15</sup>

The use of comparative methods involves extrapolation from one or more known structures to construct a new structure. It is important to consider how accurately this function can be performed in assessing whether it is a useful and worthwhile exercise. Certainly, comparisons that have been published between predicted model structures and the experimental structures determined later<sup>16-18</sup> indicate that the modeled structure is not likely to be completely accurate. This paper reports the methods that we have developed to perform comparative modeling and the criteria that we use to estimate the reliability of various parts of the structure and especially to identify potential problem areas that are likely to be particularly difficult to predict correctly.

To describe properly and illustrate the modeling techniques, we will use the mammalian serine protease family.<sup>19</sup> Members of this family are ubiquitous in nature; they are present all along the evolutionary pathway from bacteria to humans. They play important roles in a wide variety of body functions, including blood coagulation, fibrinolysis, complement activation, fertilization, and digestion. Therefore, modeling the various members of this family can provide valuable new information about diverse critical biological functions of the body.

Comparative modeling works best when there are several experimental structures and known sequences. Experimental structures for seven differ-

Received September 12, 1989; accepted January 17, 1990.  
Address reprint requests to Dr. Jonathan Greer, D-47E, Pharmaceutical Products Division, Abbott Laboratories, 1 Abbott Park Road, Abbott Park, IL 60064.

**TABLE I. Experimentally Known Three-Dimensional Structures of Serine Proteases**

Protein	Source	Resolution (Å)	Reference
Chymotrypsin	Bovine	1.8	24*
Trypsin	Bovine	1.8	25*
Elastase	Porcine	1.8	26
Kallikrein	Porcine	1.8	27
Mast cell protease	Rat	1.8	28
<i>S. griseus</i> trypsin	<i>S. griseus</i>	1.7	23
Tonin	Rat	1.8	29

\*Structures have been determined by several groups for each of these proteins. The reference shown corresponds to the coordinates used in this work.

ent serine proteases\* that can be considered members of this mammalian family have been reported in the literature and deposited in the Brookhaven Protein Structure Database.<sup>8</sup> These include chymotrypsin,<sup>24</sup> trypsin,<sup>25</sup> elastase,<sup>26</sup> kallikrein,<sup>27</sup> rat mast cell,<sup>28</sup> *Streptomyces griseus* trypsin-like protein,<sup>23</sup> and tonin<sup>29</sup> (see Table I). Several of these have been solved independently in more than one laboratory either in the same form or in alternate forms of the molecule.<sup>30-32</sup> In addition, structures have been reported for the precursor, zymogen, form of chymotrypsin<sup>33,34</sup> and trypsin.<sup>35,36</sup> To complement these structures, amino acid sequences for over 35 different serine proteases have been reported (Table II), not counting species variations. Thus this family presents an ideal system for developing comparative modeling methods and at the same time applying them to proteins involved in important and interesting biological functions.

#### COMPARATIVE MODELING METHOD: THE "SPARE PARTS" ALGORITHM

Comparative modeling requires extrapolation from known structures to produce a model of an unknown structure. Our ability to extrapolate accurately is, unfortunately, greatly limited by our still rudimentary knowledge and understanding of protein structure and energetics. Consequently, the techniques that are used to extrapolate influence the resultant model critically and may introduce considerable error. The methods that we have developed attempt to systematize the modeling process by combining all the known structures and sequences to help improve the accuracy of the extrapolation.

\*Structures have also been reported for several serine proteases from bacterial sources.<sup>20-22</sup> Although these are certainly members of this family, they differ sufficiently in their structures that comparative modeling methods are not developed enough to permit modeling them from the mammalian structures without serious errors.<sup>23</sup> Therefore, they are not considered in this work.

#### Analysis of the Structural Properties of the Family

The first steps in the modeling process are illustrated schematically in Figure 1. Let us assume that we have experimentally determined three-dimensional structures for several members of the homologous family of interest, e.g., structures "A," "B," and "C" in Figure 1a. The known structures are superimposed in three dimensions to obtain a maximal overlap of the structures (Fig. 1b). Performing this superposition of the structures is sometimes not a trivial procedure. This is true because we want to overlap the molecules based on the parts that are the same and ignore the parts that are different. The more different the molecules are from each other, the harder it is to find a unique best overlap of the common features of the structures. Several programs have been written that do this analytically.<sup>37,38</sup>

Once superimposed, there are parts of the known structures that overlap very well, indicating that the structures are closely conserved in these regions (see bold areas in Fig. 1b). We call these portions "structurally conserved regions," or SCRs, and expect that they will usually remain conserved in all members of this homologous family. These regions are usually composed of the secondary structure elements, the immediate active site, and other essential structural framework residues of the molecule. Between these conserved elements are highly variable stretches that differ significantly from one member of the family to the next. These are called "variable regions," or VRs. They are almost always loops that lie on the external surface of the protein, and they contain all the additions and deletions between different protein sequences.

After the structures have been overlapped in three dimensions and parsed into SCRs and VRs, the next step is to align their amino acid sequences. For purposes of comparative modeling, the sequence alignment is done differently from the usual methods.<sup>39</sup> Instead of relying on criteria such as amino acid identity or homology, we use strictly the three-dimensional overlap as the criterion. When the  $\alpha$ -carbons of the respective residues in the overlapped protein structures occupy the same place in three-dimensional space, then the residues are corresponded in the sequence alignment (Fig. 2). Thus the alignment is primarily concerned with the SCRs, since these are the portions of the structures that are basically the same in all protein members of the family. The alignment in the VRs is often arbitrary, unless two or more structures have similar conformations in a particular VR (for an example of this see the VR in the upper left of the structures "A" (sequence C-D-F-A) and "C" (sequence C-R-Y-V) in Figs. 1 and 2).

The resultant sequence alignment is then scruti-

TABLE II. Sequences of the Serine Proteases\*

Protein	Code	Source
Chymotrypsin	CHT†	Bovine
Trypsin	TRP†	Bovine
Elastase	ELA†	Porcine
Kallikrein	KAL†	Porcine
Mast cell protease	MCP†	Rat
<i>S. griseus</i> trypsin	SGT†	<i>S. griseus</i>
Tonin	TON†	Rat
Haptoglobin heavy chain	HPH	Human
Protein Z	PRZ	Bovine
Protein C	PRC	Human
Nerve growth factor $\alpha$ chain	NGA	Human
Nerve growth factor $\gamma$ chain	NGG	Human
Blood clotting factor VII	VII	Human
Blood clotting factor IX	FIX	Human
Blood clotting factor X	FAX	Human
Blood clotting factor XI	FXI	Human
Blood clotting factor XII	XII	Human
Plasmin	PLM	Human
Apolipoprotein A	ALP	Human
Tissue plasminogen activator	TPA	Human
Urokinase	UKH	Human
Thrombin	THR	Human
Complement factor B	CFB	Human
Complement factor 2	CF2	Human
Complement factor D	CFD	Human
Complement factor 1R	C1R	Human
Complement factor 1S	C1S	Human
Adipocyte protease	ASP	Mouse
Cathepsin G	CAG	Human
T-cell serine protease	TCL	Mouse
Hannuka factor	HAF	Human
Cytotoxic T lymphocyte protease	CTL	Mouse
Batroxobin	BTX	Snake
EGF binding protein	EBP	Mouse
<i>Drosophila</i> snake locus	DSN	<i>Drosophila</i>

\*The sequences are taken from the sequence database.<sup>8</sup>

†For those proteins for which a three-dimensional structure is available, the sequence is taken from the Brookhaven structure database.<sup>9</sup>

nized carefully to identify strong stretches or patterns of sequence homology that are characteristic of each SCR in the structure. For example, the sequence "L-S/T-V- $\pi$ -I- $\pi$ ," where  $\pi$  is a charged or polar residue, is conserved in the second SCR in Figure 2. Similarly, "G-I-A" is found in all members at the third SCR. However, in the fourth SCR, the situation is less clear. A proline usually appears in the second position but may be replaced by a glycine. Most SCRs can be identified by a characteristic (not necessarily contiguous) sequence pattern. In some cases, there is a pattern of hydrophobic or hydrophilic residues rather than specific side chains. Careful analysis of the sequences and the corresponding three-dimensional structures will usually allow some pattern to be discerned.

The identification of these sequence homology patterns is a crucial step in the modeling. As is shown below, it is essential for the correct alignment of a "new" sequence. If the characteristic sequence pattern is not present, then it is not clear how to align the "new" sequence and, consequently, how to

construct the model structure. This is one of the reasons that the comparative modeling method currently depends so heavily on the retention of sequence homology. The above steps are performed once for a protein family and need be reexamined only as new experimental structures and their sequences become available.

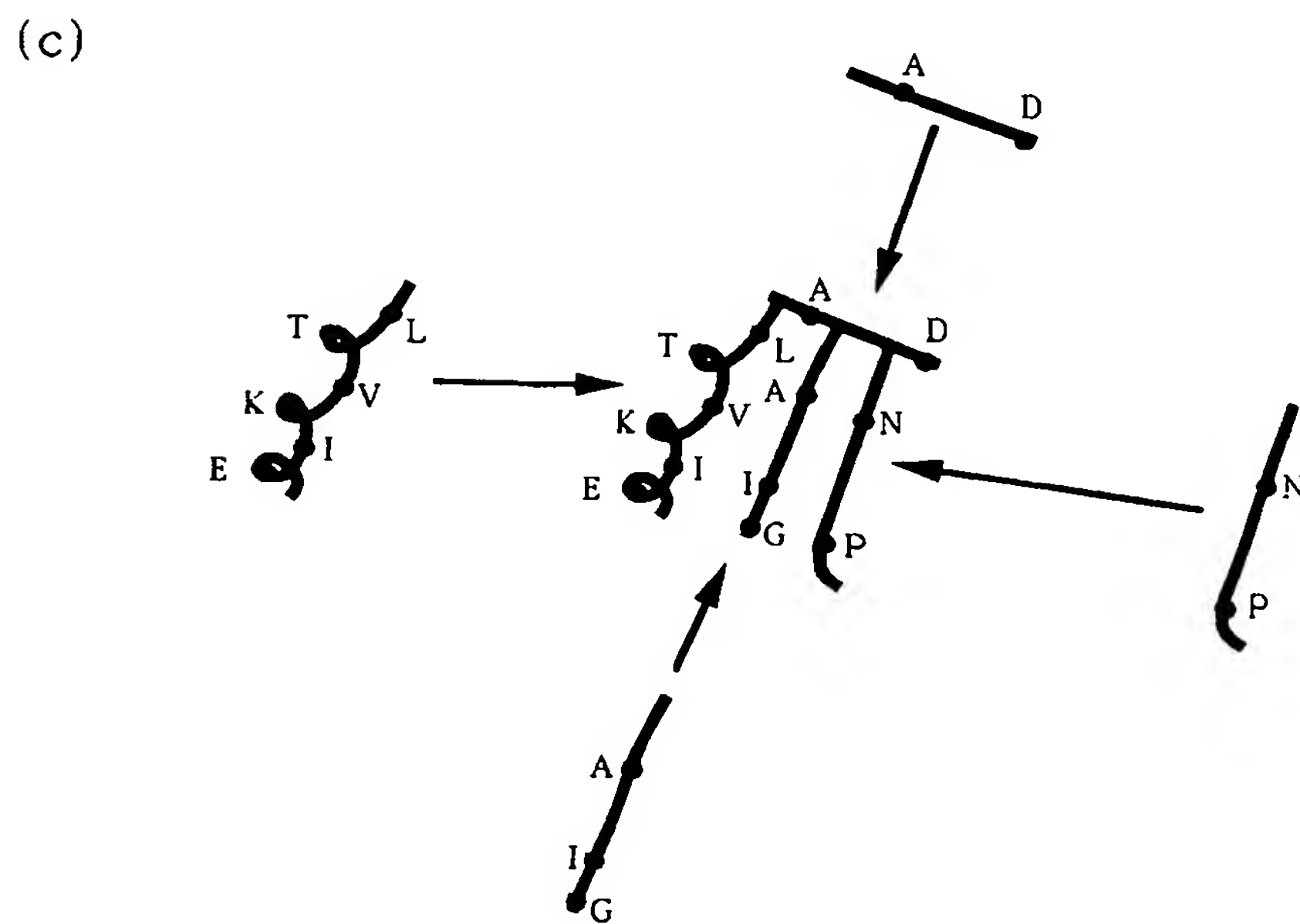
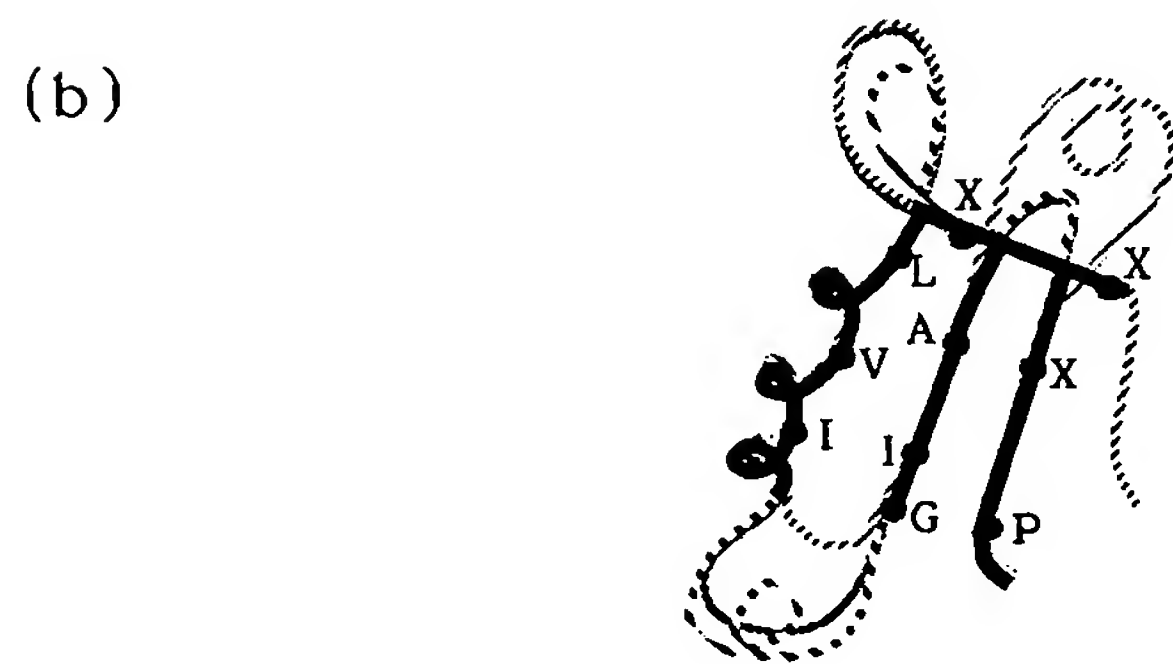
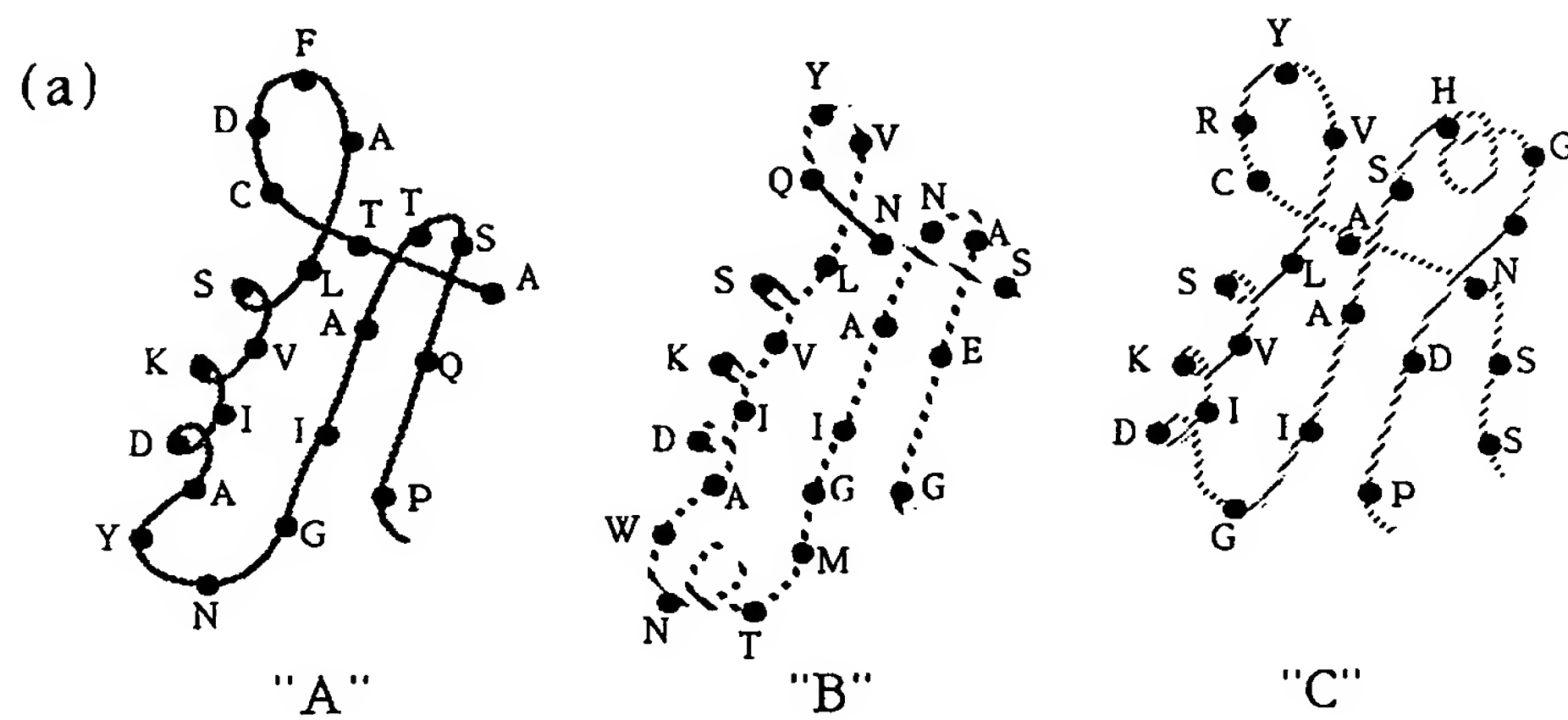
### Construction of a "New" Structure

We are now ready to begin modeling a "new" protein sequence of interest that is a clear member of this homologous family. The first step is to align this "new" sequence to the sequences of the SCRs using the previously identified characteristic sequence homology patterns. It should be possible to align an appropriate portion of the "new" sequence to each SCR with these patterns. Thus, the "new" sequence L-T-V-K-I-E fits the pattern L-S/T-V- $\pi$ -I- $\pi$  described above and can be aligned to the SCR at positions 7–12 in Figure 2 and similarly for the G-I-A sequence of the SCR at 16–18. There will occasionally be more than one possible alignment for a particular SCR, especially if the sequence pattern is a weak one. Such an example might occur in the SCR corresponding to residues 1–2 (Fig. 2). In such cases, the different possible alignments would have to be considered. As each portion of the "new" sequence is aligned to an SCR, the rest of the positions in that SCR are filled with the adjacent "new" sequence without permitting any additions or deletions within the SCR.<sup>†</sup>

The remaining residues make up the VRs. The "new" sequence in each VR is examined to see if it corresponds in length and residue character to one of the VRs of the known structures. For example, the third VR in the "new" protein in Figures 1 and 2, residues 13–15, has the same sequence length and character as this VR in protein "B" (A-W-D-S-L in "new" vs. A-W-N-T-M in "B") and therefore has been aligned to it. On the other hand, none of the known structures has a sequence that corresponds to residues 3–6 of the "new" protein in residue length.

Once the "new" sequence is aligned, as shown in Figure 2, the model building can begin. For the SCRs, usually the main chain coordinates from any one of the known structures can be taken (Fig. 1c). The side chains are mutated to those of the "new" sequence wherever necessary. In our implementation, the  $\chi_1$  angle for the "new" side chain is chosen to maximize overlap of this side chain on the old one. Further side chain  $\chi$  angles are not fitted at this

<sup>†</sup>An exception would occur if there are too few residues between this SCR and the neighboring SCR. In that case, the end residues of an SCR would be left with a deletion. For examples, see position 96 in PLM and ALP in Figure 4.



**Figs. 1a–c. Legend appears on page 321.**



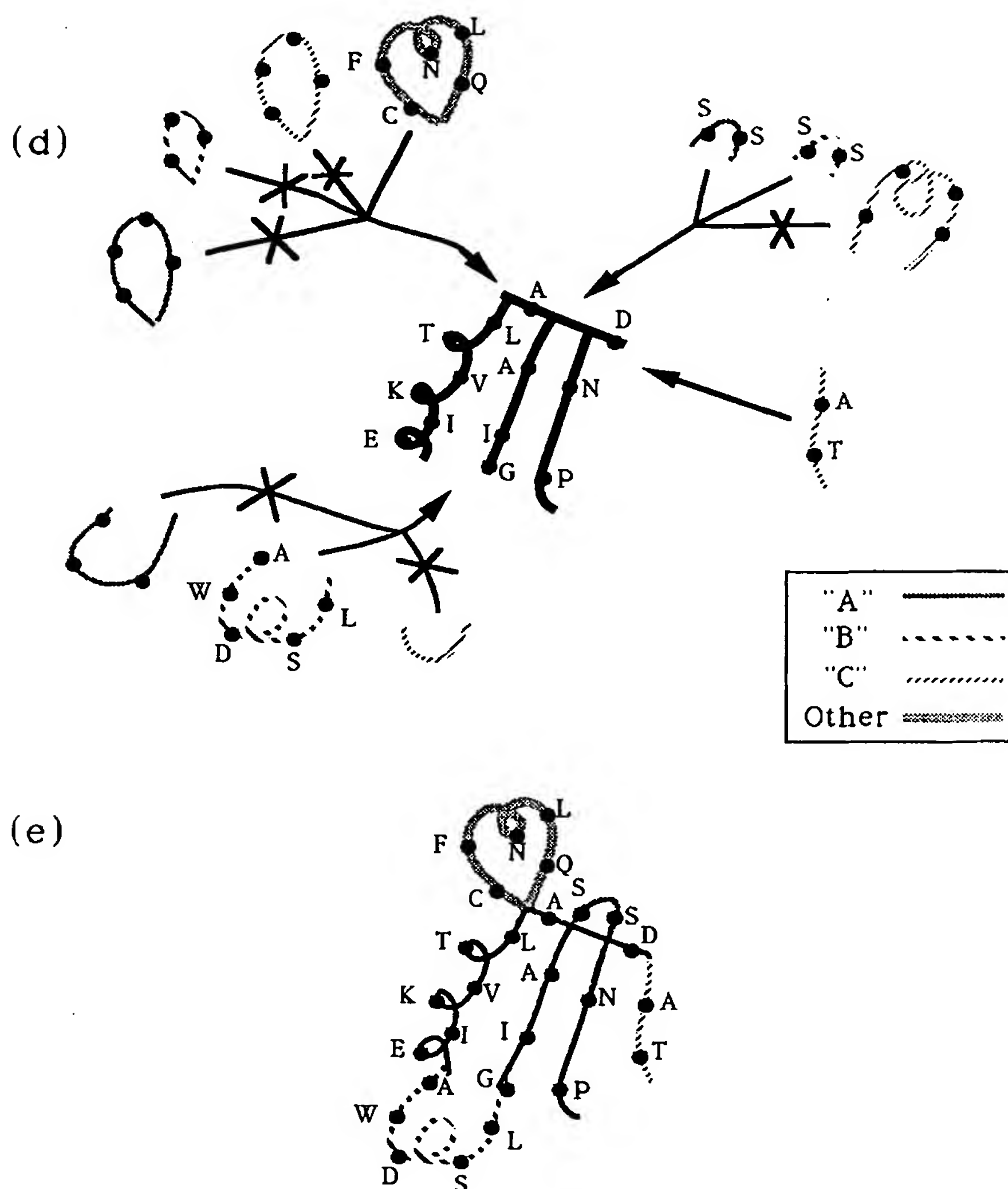


Fig. 1. Schematic representation of the comparative modeling method. **a:** There are three proteins in this homologous family whose structures are known, A, B, and C. Each protein is represented with a characteristic dashed or dotted pattern throughout the figure (see key in d). **b:** The three proteins are superimposed showing that parts of the structure are conserved (SCRs in bold lines) and parts are variable from one protein to the next (VRs, respective dashed and dotted lines). **c:** Steps in the construction of the schematic "new" model structure are presented in this and the next two parts. The SCR (bold lines) are constructed from the main chain coordinates of any one of the known structures, since they are all very close to the same. The side chains are mutated

to the "new" sequence where necessary. **d:** The various VR conformations found in the known structures are considered for each VR of the "new" protein. (Compare the respective VRs shown here with those in a and b.) The ones that do not fit are rejected, as shown by the crossed arrows. The most suitable is selected in each case. In some cases, other conformational search<sup>41-45</sup> or energetics<sup>46,47</sup> methods must be employed, since no suitable conformation can be found for that VR among the known structures (see VR in upper left corner with the sequence C-F-N-L-Q for an example in which a different, new conformation is necessary). **e:** The composite structure showing the source of the respective "spare parts" selected for the model structure.

#:	5	10	15	20
A:	ATCD-FA	LSVKID	AYN--	GIA T--SQP
B:	SNQ--VV	LTVRID	AWN TM	GIAN--AEG
C:	SSNACR-YV	LSVEIK	----	GIA SHGTDP
NEW:	TADACFNLQ	LTVKIE	AWDSL	GIA S--SNP

Fig. 2. Sequence alignment for the set of proteins in the schematic homologous family corresponding to Figure 1. The alignment is performed based solely on the overlap of the three-dimensional structures and not by sequence alignment methods. The boxes delineate the SCR as determined from the three-dimensional structure overlap. The sequence of a "new" protein is aligned based on the characteristic patterns of sequence homology found for the known structures and their sequences (see text). The single letter amino acid codes are given in Figure 4.

stage, because the nature of the different side chains usually diverges beyond this point. An automatic angle-scanning energy-optimizing routine is usually not employed, because we find that it often selects an unsuitable conformation based on minor or insignificant differences in energy. However, incorporation of recently compiled and reported rotamer libraries<sup>40</sup> may improve our ability to automate the side chain conformation selection process.

For those VRs for which an appropriate example structure can be selected from among the known structures during the alignment process described above, the main chain fragment is taken directly from the VR of that known structure, and again the side chains are mutated to fit the "new" sequence (see examples in Fig. 1d and e). Remaining VRs, such as the VR at positions 3–6 in Figures 1 and 2, have no good example to build upon among the known structures. Therefore, they have to be constructed by more complex conformation search methods<sup>41–45</sup> and energetic considerations.<sup>46,47</sup> We have also used the protein database as a source of tentative starting conformations for loops of this type. The two ends of the particular loop that lie in the adjacent SCRs are taken, and the Brookhaven Protein Structure Database<sup>8</sup> is searched for structural fragments that closely match the conformation of these ends (specifically, the  $\alpha$ -carbon positions) and have the appropriate number of residues between the ends. Thus, for example, in the VR at positions 3–6 of the "new" sequence, the structures of the two ends corresponding to residues "A-D-A" (positions –1 to 2) and "L-T-V- . . ." (positions 7 to 9. . .) are selected and typically ten structures that match these two ends best [lowest root mean square (rms) deviations] and that have five residues in between are chosen and examined. Obviously, any such conformation that would collide with the remainder of the constructed "new" structure is eliminated. Similarly, if the conformation does not pack well or buries a single charged group, it is rejected. In this way, one or a small number of initial conformations can be selected for this loop. Clearly this latter method does not produce an exhaustive list; however, it does provide loop conformations that are known to appear in other proteins. This method has previously been described by Kraulis and Jones<sup>48</sup> for constructing protein structures from fragments using nuclear magnetic resonance (NMR) data or using crystallographically generated electron density maps.

It is important to emphasize that it is the initial spatial alignment of the three-dimensional structures performed above (Fig. 1b) that allows this convenient clipping of fragments or "spare parts" from the various known structures to construct a composite model structure of the "new" protein. Thus overlapping of the structures is essential for two crucial steps: for the original alignment of the sequences

and for clipping together fragments in the subsequent construction of each "new" model protein structure. Note that, the more experimentally known structures available, the more likely the boundaries of the SCRs will be well-defined and the greater the likelihood that a suitable example known structure, i.e., spare part, will be found for each "new" VR among the known structures.

Because the different portions of the "new" constructed model structure arise from quite different sources (Fig. 1e), we can assign appropriate qualitative reliability confidence levels to the respective parts of the structure. Clearly, the conformations in the SCRs are the most reliable, especially when several known structures are available and have been compared in detail. This is true only if the respective SCRs of the "new" sequence retain the characteristic homologous sequence patterns in those SCRs. When they do not, and this is illustrated below, it should be regarded as a red flag warning that something different may be happening in this region of the particular "new" protein. In such cases, great care must be taken in constructing the "new" structure and the confidence level in this portion of the molecule reduced appropriately. For the VRs, when a good model structure appears for the respective VR among the known structures, then a fair confidence level can be assigned. This level is not as high as in the SCRs because of the greater inherent variability of these loop regions. The actual confidence level for each such VR would depend on the size of the loop (larger loops have more degrees of freedom and are therefore less reliably determined), how good the sequence homology is to the known structure selected to model this VR, and how well the chosen conformation fits and packs onto the rest of the "new" structure. The confidence is lowest for those VRs with no good model loop among the known structures. When conformational methods or database search techniques have to be employed, experience<sup>16–18</sup> shows that we are not able to predict the conformation reliably in many of these cases.

The resultant structure is then examined in detail to identify possible serious errors. For example, the inner cores of the structure are checked to be sure that no inappropriate charges have been buried. The modeling method makes it unlikely that there are serious steric contacts between main chain atoms in the model. This is because the main chain coordinates in the SCRs were taken from the known structures. Only if the bad contact was present in the parent known structures will it be found in the model. In the VRs, absence of main chain overlap was a major criterion for acceptance of a conformation. Therefore, in practice, bad steric contacts can usually be relieved by side chain rotations; occasionally by having to select an alternative conformation for a VR. The model structure can then be introduced into an energy-minimization program<sup>46,47</sup> to

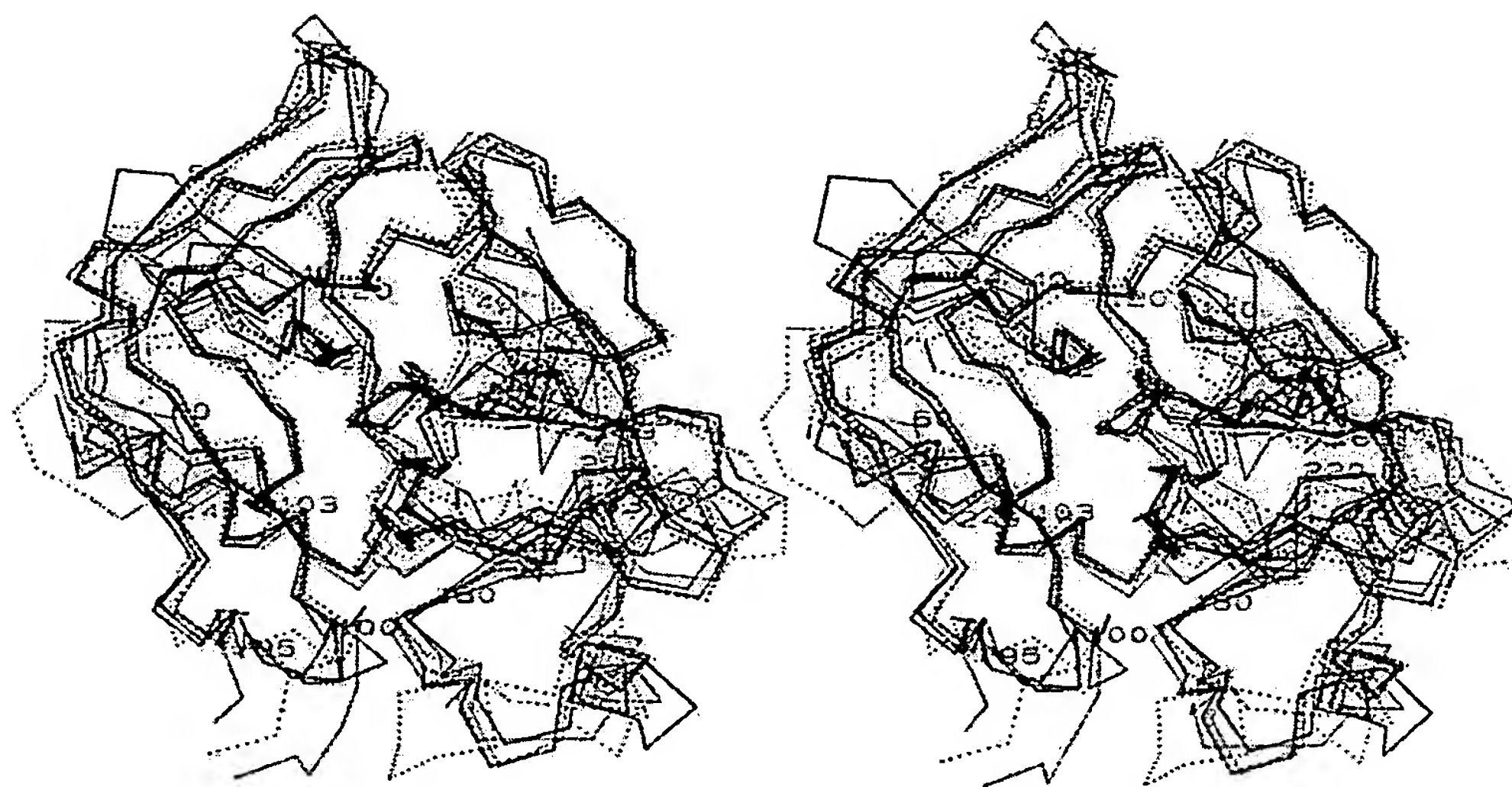


Fig. 3.  $\alpha$ -Carbon plots showing the superposition of the seven experimentally known serine protease structures (see Table I). Note the SCRs, where the seven structures are virtually the same, and the VRs, where the structural divergence is considerable.

The key to the plots is as follows: CHT, green; TRP, solid red; ELA, solid blue; KAL, solid cyan; MCP, dotted green; TON, dotted red; SGT, dotted blue.

relieve any remaining minor steric contacts and to optimize the bond angles and torsional angles. Typically, we begin with restrained or template-forced<sup>49</sup> minimization, forcing the atomic coordinates to remain close to their initial positions. This allows bad contacts to relax without introducing large artificial distortions into the structure. After several hundred cycles, the constraints can be progressively removed and the energy of the molecule minimized.

Experience with comparative molecular modeling<sup>16-18</sup> has shown that these models are never completely accurate, especially in the conformations selected for the VRs. Therefore, any uses of the model structure and predictions that are made based on it should take into consideration this fact and the local reliability confidence levels discussed above. We always regard such a model as a working hypothesis; as new data emerge, whether from structural, spectroscopic, or biochemical sources, the model is modified, corrected, improved, and refined to fit the new data.

## RESULTS AND DISCUSSION

### Initial Alignment of Known Serine Protease Structures

Our analysis of the serine proteases began some years ago with the superposition of the then three known structures of chymotrypsin, trypsin, and elastase.<sup>3,4</sup> Since then, coordinates for four more

structures have appeared in the literature and data bank<sup>8</sup>: kallikrein,<sup>27</sup> mast cell protease,<sup>28</sup> *Streptomyces griseus* trypsin-like protease,<sup>23</sup> and tonin.<sup>29</sup> The overlap of these additional structures was performed by first superimposing the added experimental structure onto the rest using the  $\alpha$ -carbon positions of residues 195, 57, and 102.<sup>‡</sup> The superposition was refined, when necessary, using repeated cycles of least squares fit on related residue  $\alpha$ -carbon positions between the protein to be fit and chymotrypsin. Only residues whose positions were conserved in the two structures were fitted in the calculation; that is, any residue was eliminated from the fitting if the deviation in  $\alpha$ -carbon positions was more than 2 Å on the first cycle and then more than 1.5 Å on subsequent cycles. Figure 3 shows the final overlap obtained for these seven structures. Virtually all the SCR positions identified in the previous analysis of the first three proteins<sup>4</sup> remain structurally conserved. The repertoire of conformations, i.e. "spare parts," found for each of the VRs is increased with the larger number of structures, as expected (Table III).

<sup>‡</sup>The residue numbering given for the serine proteases follows that of the chymotrypsinogen molecule throughout this paper, in the text, in the tables, and in the figures.



TABLE III. Sequence Lengths of the Different VRs in the Serine Proteases

VR	VR									
	35-41	59-62	72-80	97-101	125-133	146-151	166-179	185-189	203-206	217-224
CHT	7	4	9	5	9	4	14	3	4	8
TRP	5	3	9	5	7	6	14	5	0	8
ELA	10	5	9	7	9	5	16	4	4	10
KAL	6	3	9	9	7	8	14	5	0	9
MCP	10	3	9	5	9	5	13	5	0	6
SGT	2	8	7	3	7	5	15	6	5	8
TON	4	3	9	16	7	6	14	5	0	9
HPH	6	16	0	1	8	4	31	5	6	7
PRZ	7	1	4	5	11	2	13	3	4	6
PRC	7	5	9	5	13	10	14	5	4	8
NGA	6	3	9	16	7	6	14	5	0	9
NGG	6	3	9	12	7	6	14	5	0	9
VII	6	8	9	5	12	5	19	5	4	8
FIX	6	5	9	7	11	5	14	5	4	8
FAX	7	5	8	5	11	5	14	5	4	8
FXI	9	8	9	5	9	5	15	5	3	8
XII	5	8	9	5	9	6	16	5	7	8
PLM	7	8	9	-1	9	4	16	5	4	8
ALP	7	8	9	-1	9	4	7	5	4	8
TPA	12	8	9	5	9	6	16	11	4	8
UKH	11	8	9	7	9	6	16	5	4	8
THR	8	13	10	6	12	11	14	8	6	8
CFB	10	8	1	14	25	3	27	7	4	14
CF2	6	8	5	14	25	0	28	4	4	21
CFD	6	7	10	5	11	5	10	3	0	9
C1R	4	12	8	8	9	3	19	5	6	6
C1S	3	5	10	12	11	4	23	4	7	5
ASP	6	8	9	5	11	5	16	3	0	9
CAG	9	3	9	5	9	4	14	5	0	6
TCL	6	4	8	5	9	5	18	5	0	9
HAF	6	4	9	5	9	5	18	5	0	10
CTL	9	3	9	5	9	5	15	5	0	6
BTX	6	3	13	2	7	6	13	5	0	9
EBP	6	3	9	16	7	6	14	5	0	9
DSN	15	5	10	5	6	5	20	4	7	8
Min.	2	1	0	-1	7	0	7	3	0	5
Max.	15	16	13	16	25	11	31	11	7	21
Range	13	15	13	17	18	11	24	8	7	16

Based on the above superposition of structures, the seven sequences were aligned as described above (Fig. 4). Note the wide variation in the length of the different sequences in each of the VRs (Table III). Using this sequence alignment, we attempted to identify the critical conserved stretches of amino acid sequence. These are characteristic of each SCR and are essential for aligning "new" sequences to the known ones, which is the first step in modeling. The patterns that were chosen are noted in Figure 4 at the bottom of the sequence list in the row labeled "CON." In some cases, there was almost complete conservation of a single amino acid. In other cases, only a hydrophobic side chain (typical of an internal position) seems to be required.

Because the identification of these stretches is such an important part of the modeling process, it is worthwhile describing in more detail how this is performed. In most cases, it is straightforward; the conserved sequence is readily apparent. For example, the characteristic I/V-I/V-G-G at positions 16-

19 (see Fig. 4) defines the first SCR very clearly. Similarly, the remarkably conserved S/T/A-G-W-G sequence of residues 139-142 is highly characteristic of this SCR. On the other hand, the SCRs at 63-71 and 81-96 have no such clearly conserved sequence, and it is often very difficult to align a "new" sequence to the known ones in these SCRs unambiguously.

To understand better why this latter group of SCRs has so little sequence homology, the structures of these two regions were examined in detail (Fig. 5). It is immediately apparent that these SCRs lie on the surface of the molecule, and virtually all the residues point out into solvent and thus can and do vary with little consequent effect on the rest of the structure. However, in each SCR, there are some side chains that point into the interior of the molecule and thus are more conserved. In the SCR at 63-71, this includes the Val (or comparable aliphatic residue) at 66 and the aliphatic side chain at 68. The Gly at position 69 appears to be almost com-



CNO	16	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	CNO
CHT	IVNCEAVPGSWPQVSLQDKT	-----GFHFCGSLINENWVVTAAHCGVT	-----TSDVVVAGEFDQGS	-----SSE-KIQKLIKAKVFNKSYNSLTI	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	CHT
TRP	IVGCTCGANTVPYQVSLNS	-----GYHFCGSLINSQWVSAACHYKS	-----GQVRLGQDNINV	-----VEGNQOFISAKSIVHPSYNSNTL	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	TRP
ELA	VVGTEAQRNSWPSQISLQYRS	-----GSSWAHTCGGLIRQNWMVMTAAHCVDR	-----LTFRVWVGEHNLN	-----NNGTEQYVGVQKIVVHPYWNDDVA	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	ELA
KAL	IIGCEKEKNSHPWQVAYHY	-----SSFQCGVLVNPKWVLTAAHCKN	-----DNYEVLGRHNLFE	-----NENTAQFFGVTAQFFPHPGFNLSADCK	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	KAL
MCP	IIGVESIPHSRPMALHDIVT	-----EKGLRVICGGFLISRFVLTAAHCKG	-----REITVILGAHDVRK	-----AESTQKIKVEKQIIHESYNSVPN	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	MCP
SGT	VVGTRAAQGEFFPMVRLS	-----MCGGALYAQDILVLTAAHCVSGS	-----GNNTSITATGGVVDLQS	-----GAAVKVRSTKVLQAPGYNGT	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	SGT
TON	IVGGYKCEKNSOPWOAVIN	-----EYLGGVLIDPSWVITAACHCYSN	-----NYOVLLGRNNLFK	-----DEPFAORRLVROSEFRHPDYIPLIVTNDTEOPY	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	TON
HPH	ILGGLDAKGSFPWQAKMVSH	-----HNLTGTATLINEQWLLTTAKNLF LNHSENATAKDIAPTLTLYVGKK	-----QLVEIEKVVLHPNYSQV	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	HPH
PRZ	RPQSQQNLPLFPWQVKTNSE	-----GKDFCGGVLIQDNFVLTATCS	-----LLYANISVKT	-----RSHFRLHVRGVHVHTRFEADTG	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	PRZ
PRC	LIDKMTTRRGDSPWQVLLDSK	-----KKLACGAVLIHPSWVLTAAHCMDES	-----KKLLVRLGEYDLRR	-----WEKWELDLDIKEVFVHPNYSKSTT	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	PRC
NGA	VQSQVDC=NSQPHVAVYRF	-----NKYQCGGVLDRNWLTAACHCYN	-----KYQVWLGNKNNFL	-----DEPSDQHRLLVSKAIPHDPFNMSLLNEHTPQPE	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	NGA
NGG	IVGGFKCEKNSQPHVAVYRY	-----TQYLCGGVLLDPNWLTAACHCWD	-----NYKVWLGNKNNFL	-----TEPSAQHRFVSKAIPHDPFNMSLMRFL	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	NGG
VII	IVGGKVCPCGPCWQVLLLVN	-----GAQLCGGTLINTIWWVSAACHCFDKI	-----KNWRNLIAVLGEHDLSE	-----HDGDEQSRRAQVVIIPSTYVPGTT	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	VII
FIX	VVGEDAERGQFPWQVLLHGE	-----IAAFCGGSIVNEKVVTAACHIKPG	-----VKITVAGEHNTK	-----PEPTEQKRNVIRAIPIYHSYNASINK	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	FIX
FAX	IVGGRDCAEGECPWQALLVNEE	-----NEGFCGGTILNEFYVLTAAHCLHQA	-----KRFTRVVGDRNTQE	-----GDEEMAHEVENTVKHSRKFVKEY	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	FAX
FXI	IVGGTASVRGEWPWQVTLHTSP	-----TQRHLCCGSIIGNQWILTAACHCFYGV	-----ESPKILRVYSGILNQSE	-----IKEDTSFFGVQEI IHDQYKMAES	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	FXI
XII	VVGGLVALRGAPYIAALYW	-----GHSFCAGSLIAPCWVLTAAHCLQDR	-----PAPEDLTVVLGQERRNH	-----SCEPCQTLAVRSYRLHEAFSPVSY	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	XII
PLM	VVGCVVAHPHSPWQVSLTRF	-----GMHFCGGTILISPEWVLTAAHCLKES	-----PRPSSYKVLGAHQEVN	-----LEPHVQEIIEVSRLFLEPTRK	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	PLM
ALP	IVGGCVVAHPHSPWQVSLTRF	-----GKHFCGGTILISPEWVLTAAHCLKKS	-----SRPSSYKVLGAHQEVN	-----LESHVQEIIEVSRLFLEPTQA	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	ALP
TPA	IKGGLFADIASHPWQAAIFAKHR	-----RSPGERFLCGGILISSCWILSAACHCFQER	-----FPPHLLTVILGRTYRVV	-----PGEEEQKFEVEKYIIVHKEFDDDTY	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	TPA
UKH	IIGGEFTTIENQPFWAAIYRRHR	-----GGSVTVYVCGGSLISPCWVISATHCFIDY	-----PKKEDYIYVIGRSRLNS	-----NTQGEKMFVENLILHKDYSADTL	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	UKH
THR	IVEQDAEVLSPWQVMLFRKS	-----PQELLCGASLISDRWVLTAAHCLLYPPW	-----BKNTVDLLVRIGKHSRTR	-----YERKVEKISMLDKIYIHPRYNWKEN	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	THR
CFB	WEHRKGTDYHKQPWQAKISVIR	-----PSKGHESCMGAVSEYFVLTAAHCFITVD	-----DKEHSIKVSVGGEK	-----RDLEIEVVLVHFPNYPNNGKKEAG	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	CFB
CF2	NMSANASDQERTPWHTIKPK	-----SQETCRGALISDQWVLTAAHCFRDG	-----NDHSLWRVNVGDPKSO	-----WGKELLIEKAVISPFGDFVAKKNQ	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	CF2
CFD	ILGGEAEAEHARPYMASVQLN	-----GAHLCCGGVLVAEQWVLSAAHCLLEDA	-----ADGKVQVLLGATHLPQ	-----PEPXXXITIEVLRVAVPHDPDSQPD	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	CFD
C1R	IIGQKAKMGNFPPWQVFTNI	-----HGRGGGALLGDRWLTAAHCLLYPKEH	-----EAQSNASLDVFLGHTNVVEE	-----LMKLGHPIRRVSHPDPYRQDES	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	C1R
C1S	IIGSDADIKNFPWQVFFD	-----NPWAGGALINEYVLTAAHCVVEGN	-----REPXYVVGSTSVQT	-----SRLAKSKMLTPEHVF IHPGWKLLLEVPEG	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	C1S
ASP	ILGQEAHAHARPYMASVQVN	-----GTHVCGGTLLEQWVLSAAHCLMDGV	-----TDDDSVQVLLGAHLSA	-----PEPYKRWYDVQSVVPHPGSRPDSL	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	ASP
CAG	IIGRESRPHSRPYMAYLIQSP	-----AQOSRCGGFLVREDFVLTAAHCVGNS	-----INVTLGAHNIQR	-----RENTQQHITARRAIRHPQYNQRTI	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	CAG
TCL	IIGDTPVPHSRPYMALLLKS	-----SNTICAGALIEKNWVLTAAHCVNG	-----KRSKFI LGAHSINK	-----EPEQOILT VKKAFPPCYDEYTR	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	TCL
HAF	IIGNEVTPHSPYMWLLSLD	-----RKTICAGALIAKDWVLTAAHCLN	-----KRSQVILGAHSITR	-----EEPTKQIMLVKKEFPYPCYDPATR	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	HAF
CTL	IIGHEVPHSRPYMALLSIK	-----DQQPEAICGGFLIREDFVLTAAHCEGSI	-----INVTLGAHNIKE	-----QEKTOQVIVPVKICPHDPYNPKTF	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	CTL
BTX	VIGDECDINEHPFLAFMYYS	-----PRYFCGMTLINEWVLTAAHCLNRRF	-----MRIHLGKHAGSVANYDEVVRYPKKFCIPNKKKNVITD	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	BTX
EBP	VVGGFNCEKNSQPHWQVAVYQ	-----KEHICGGVLLDRNWLTAACHCYVDQ	-----YEVWLGNKMLFQ	-----EEPSAQHRLVSKSFPHPGFNMSLLMLQTPPG	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	EBP
DSN	IVGGTPTRHGLFPHMAALGWTOGSGSKDQDIKWGCGGALVSELYVLTAAHCAATSG	-----ANHRTWFAWRPQLN	-----ETSATQDDIKILIIIVLHPKYRSSAY	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	DSN
CON	i1GG	Pwqv λ	CGG li	wVLTAAHC	v λG	λ λ π λ hp y	80 78	?	?									CON
S-S	157	22	58	111	42													S-S

Fig. 4. Part 1. Legend appears on page 328.

CNO 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 CNO

CHT ---NNDITLLKLS---TAASFQTSVAVCLP-SASDD---FAAGTTCVTTGWGLTRY---ANTPDRLOQASPLLSNT-NCKK---CHT

TRP ---NNDIMLIKLS---SAASLSRVASISLP-T--SC---ASAGTQCLISGWNTKSSG---TSYPDVLKCLKAPILSNS-SCKS---TRP

ELA ---AGYDIALRLA---QSVTLNSYVQLGVLPRA-GTI---LANNSPCYITGWGLTRTN---GOLAQTLQAYLPVTDYA-ICSSS---ELA

KAL ---DYSHDLMLRLQ---SPAKITDAVKVLELP-T--QE---PELGSTCOASGWGSIIEPGPD---DFEFPDELQCVQLTLQNT-FCAB---KAL

MCP ---LHDIMLLKLE---KKVELTPAVNVVPLP-SPSOF---IHGAMCWAAGWKGTGVR---DPTSYTLREVELRINDEK-ACVD---MCP

SGT ---GKDWALIKLA---QPIN---OPTLKIA-T-TTA---YNQ-GTFTVAGWGANREG---GSQORYLLKANVPFVSDA-ACRS---SGT

TON HDHSNDIMLLHLS---EPADITGGVKVIDLP-T--KE---PKVGSICLASGWGSTNPSE---MVVSHDLQCVNIHLISNE-KCIE---TON

HPH =====DIGLIKLS=====QKVSVNERVMPICLP=SK=DY=====AEVGRVGYVSGWGRNAN=====FKFTDHLKYVMLPVADQD=QCIRHYEGSTVPEKKTPKS

PR2 =====HNDVALLDLA=====RPVRCPDAGRVPVCTADADFAD=====SVLLPQPGVLGGWTLRGR=====EMVPLRLRVTHVEPAE==CGR=====

PRC =====DNDIALHLA=====QPATLSQTIIVPICLPDSGLAER=====ELNQAQOETLVGTWGYHSSREKE=AKRNRFTVLFNFIKIPVVPHN=ECSE=====

NGA DDYSNDIMLLRLS=====KPADITDVVKPITLP=T==EE=====PKLGSTCLASGWGSTPIK=====FKYPDDDLQCVNLKLLPNE=DCDK=====

NGG YDYSNDIMLLRLS=====KPADITDVVKPITLP=T==EE=====PKLGSTCLASGWGSTIPTK=====FOFTDDLYCVNLKLLPNE=DCAK=====

VII =====NHDIALRLH=====QPWLTDHWPVLCLPERTFSE=====RTLAFVRFSLVSGWGLLDR=====CATALELMVNLVPRIMTQ=DCLOQ=====SR

FIX =====YSHDIALLELD=====EPLEINSYVTPICLADRDYTN=====IFSKFGYGYVSGWKVFN=====GRSASILQYLKVPVLDRA=TCLE=====

FAX =====DFDIAVRLK=====TPIRFR=NVAPACLP=PEKDWA=====ETLQTKTIGVSGFGRTHEK=====GRLSSTLKMLEVPYVDRS=TCLE=====

FXI =====GYDIALKLE=====TTVNYTDSQRPICLP=SKGDR=====NVYTDWVTGWYRKL=====DKIQTLOKAKIPLVNNE=ECQK=====

XII =====QHDIALRLQEDADGSCALLSPYQVCLP=SGAAR=====PSETTLQVAGWGHQFEGA=====EEYASFLQEAQVFLSLE=RCSAP=====

PLM =====DIALKLS=====SPAVITDKVIPACLP=SPNYV=====VADRTECFITGWGETQG=====TFGAGLLKEAQLPVNIENK=VCNRY=====

ALP =====DIALKLS=====RPAVITDKVMPACLP=SPDYM=====VTARTECYITGWGETQG=====TFGTGLLKEAQLLVNIENE=VCNH=====

TPA =====DNDIALQLKSD=SSRCAQESSVVRTVCLP=PADLQ=====LPDWTCELSGYKHEALS=====PFYSERLKEAHVRLYPSS=RCTSQ=====

UKH =====AHNDIALKIRSK=EGRCAQPSRTIQTICLP=SMYND=====PQFGTSCEITGFKENSTD=====YLYPEQLKMTVVVKLISHR=ECQQP=====

THR =====LDRDIALKLS=====RPIELSDYIHPVCLPDKQTA=====KLLHAGFKGRVTGWGNRRRETWTTSVAEVQPSVLOVNLPLVERP=VCKA=====

CFB EFYDYDVALIKLS=====NKLKYGTIRPICLPCTEGTTRALRLPPTTTCQOQKEELPAQDIKALFVSEE=====EKLTREKVIYKNGDKKGCERDA=====QYAPGYDK

CF2 EFYGDIALKLA=====QVKMSTHARPICLPCTMEANLALRRPQGSTCRDHENELNKSQVPAHFVAL=====NGSKLININLKMGEVWTSCEVVS=====QEKTFMFPN

CFD =====IDHDLQLLS=====EKATLG=====PAVRPLPWQVRDR=====DVAPGTLCDVAGWGIIVNHA=====GRRRPSLOHVLLPVLDR=TCRL=====

C1R YNFEGLIALLELE=====NSVTLGNLLPICLP=DNDTF=====YDLGLMGYVSGFGVME=====KIAHDLRFVRLPVANPQ=ACENW=====LR

C1S TNFDNDIALVRLK=====DPVKMGPTVSPICLPCTSSDY=====NLMGDGLISGWGRAEK=====RDRAVRLKAARLPVAPLR=KCKEV=====KVEKPT

ASP =====EDDLILFKLS=====QNASLG=====PHVRPLPLQYEDK=====EVEPGTLCDVAGWGVVTHA=====GRRPDVHLQLRVSIMNRT=TCNLR=====

CAG =====QNDIMLLQLS=====RRVRNRNPNVVALP=RAQEG=====LRPGTLCVAGWGRVSM=====RRGTDTLREVQLRVQDR=QCLR=====

TCL =====EGDLQLVRLK=====KKATVNRNVAIHLHP=KKGDD=====VKPGTRCRVAGWGRFGNK=====SAPSETLREVNTIVDRK=ICNDE=====K

HAF =====EGDLKLLQLT=====EKAKINKYVTILHLHP=KKGDD=====VKPGTMQCVAGWGRTHNS=====ASWSDTLREVNTIIDRK=VCNDR=====N

CTL =====SNDIMLLKLS=====SKAKRTRAVRPLNLP=RRNVN=====VKPGDVICYVAGWGRMAPM=====GKYSNTLQEVLTQKDR=ECES=====

BTX =====KOIMLIRLD=====RPVKNSEHIALSLP=S==NP=====PSVGSVCRIMGWCAITTSE=====DTYPOVPHCANINLFNNT=VCRE=====

EBP ADFSNDIMLLRLS=====KPADITDVVKPIALP=T==KE=====PKPGSTCLASGWGSTIPTR=====WQKSDDLQCVFITLLPNE=NCAK=====

DSN =====YHDIALKLT=====RRVKFSEQVRPACLW=Q=====C=====GAPHTTVVAGWGRTEFL=====GAKSNALRQVDLDVSPQM=TCCKQI=====YRK

CON Di Ll+L 50 ? 1 ? 232 c logWG 201 L λ λ λ C 22 182

S-S

Fig. 4. Part 2. Legend appears on page 328.

CNO	175	180	185	190	195	200	205	210	215	220	225	230	235	240	245	CNO																																																																																	
CHT	-YWG	TKIK	-DAM	ICAGA	-----	SGV	SCMGDS	GGPL	CKKN	---GAW	TLVGIV	SWGSS	-TCST	S-----	TPGV	YARV	-----	TAL	VNV	QQT	LAAN																																																																												
TRP	-AYP	QIT	-SNM	FAGYL	-----	QGG	KDCQ	QD	SGP	PVCS	-----	GK	LQIV	SWG	S---CAQ	K-----	NKPG	VYTKV	-----	CNY	SVW	IKQ	TIASN																																																																										
ELA	SYW	STVK	-NSM	VCAGD	-----	G	-VR	SCQ	QD	SGG	PLHCLVN	---GQ	YAVH	GVTS	FVSR	LGCNV	T-----	RKPT	VTRV	-----	SAY	ISW	INN	VIASN																																																																									
KAL	-AHP	BKV	-ESM	LACAGYL	-----	P	GK	DT	CMG	D	SGG	PLICN	-----	GM	QGIT	SWGHT	-PCGS	A-----	NKPS	IYTKL	-----	IF	YLD	WIB	BITENP																																																																								
MCP	-YR	--	Y	EYKFQ	VCV	GSP	-----	TL	RAAF	M	D	SGG	PLICA	-----	G	VAH	GIV	SYGHP	-DAK	-----	PPA	I	FTRV	-----	ST	VPT	INAVIN																																																																						
SGT	-AYG	NEI	VAN	EEI	CAGYP	D-----	TG	VD	TCQ	D	SGG	PMFRKDN	---ADE	W	IQV	GIV	SWG	Y---CARP	-----	GYP	G	VYTEV	-----	ST	FAS	AI	ASAARTL																																																																						
TON	-TYK	DNVT	-DVM	LACAGEM	-----	EGG	DT	CAG	D	SGG	PLICD	-----	GV	LQIT	SGG	ATP	-CAKP	-----	KTP	AI	YAKL	-----	IK	FTS	WIK	KVMKENP																																																																							
HPH	PVG	VQPI	ILNE	HTFCAGMS	=====	KYQ	ED	TCY	G	DAGS	AF	AVHDL	E	ED	WTY	ATG	IL	S	F	D	K	=====	EY	G	VYV	KV	=====	TS	IQD	WVQ	K	TI	AE	N																																																															
PRZ	==A	L	N	ATV	T	R	T	S	C	E	R	G	=====	AA	A	A	A	R	W	V	A	G	=====	G	P	L	L	I	K	V	=====	P	R	Y	A	L	W	L	R	Q	V	T	O	Q	P	S	*																																																		
PRC	==V	M	S	N	M	S	E	N	M	L	C	A	G	I	L	=====	G	D	R	A	C	E	G	D	SGG	P	M	V	A	S	F	H	=====	G	L	L	=====	H	N	Y	G	V	Y	T	K	V	=====	S	R	Y	L	D	W	I	H	G	H	I	R	D	K	E	A	P	Q	K	S	W	A	P																											
NGA	==A	H	K	M	K	V	T	D	A	M	L	C	A	G	E	M	=====	D	G	S	Y	T	C	E	H	D	S	G	G	P	L	I	C	D	=====	G	I	L	Q	I	T	S	W	G	P	E	P	=====	T	E	P	S	V	Y	T	K	L	=====	I	K	F	S	S	W	I	R	E	T	M	A	N	N	P																								
NGG	==A	H	I	E	K	V	T	D	A	M	L	C	A	G	E	M	=====	D	G	K	D	T	C	K	G	D	S	G	G	P	L	I	C	D	=====	G	V	L	Q	I	T	S	W	G	H	T	P	=====	D	M	P	G	V	W	T	K	L	=====	N	K	F	T	S	W	I	K	D	T	M	A	K	N	P																								
VII	KV	G	D	S	P	N	I	T	E	Y	M	F	C	A	G	S	=====	D	G	S	K	D	S	K	G	D	S	G	G	P	H	A	T	H	Y	R	=====	G	T	W	L	T	G	I	V	S	W	G	Q	=====	C	A	T	V	=====	G	H	F	G	V	Y	T	R	V	=====	S	Q	Y	I	E	W	L	Q	K	M	R	S	E	R	P	R	G	V	L	L	R	*										
FIX	==S	T	K	F	S	I	Y	S	H	M	F	C	A	G	Y	H	=====	E	G	K	D	S	C	Q	D	S	G	G	P	H	V	T	E	V	E	=====	G	T	S	F	L	T	G	I	S	W	G	E	E	=====	C	A	M	K	=====	G	K	Y	G	I	Y	T	K	V	=====	S	R	Y	V	N	W	I	K	E	K	T	K	L	T																		
FAX	==S	S	F	T	I	T	P	N	M	F	C	A	G	Y	D	=====	T	O	P	E	D	A	C	Q	D	S	G	G	P	H	V	T	R	F	K	=====	D	T	Y	F	V	T	G	I	V	S	W	G	E	G	=====	C	A	R	K	=====	G	K	F	G	V	Y	T	K	V	=====	S	N	F	L	K	W	I	D	K	I	M	K	A	R	A	G															
FXI	=R	Y	R	G	H	K	I	T	H	K	M	I	C	A	G	Y	R	=====	E	G	K	D	A	C	K	G	D	S	G	G	P	L	S	C	K	H	N	=====	E	V	H	L	V	G	I	T	S	W	G	E	G	=====	C	A	Q	R	=====	E	R	P	G	V	Y	T	N	V	=====	V	E	Y	V	D	W	I	L	E	K	T	Q	A	V																
XII	=D	V	H	G	S	S	I	L	P	G	M	L	C	A	G	F	L	=====	E	G	T	D	A	C	Q	D	S	G	G	P	L	V	C	E	D	A	E	R	R	L	T	L	Q	I	I	S	W	G	S	G	=====	C	G	D	R	=====	N	K	P	G	V	Y	T	D	V	=====	A	Y	L	A	W	I	R	E	H	T	V	S																			
PLM	=E	F	L	N	G	R	V	Q	S	T	E	L	C	A	G	H	L	=====	A	G	T	D	S	C	Q	D	S	G	G	P	L	V	C	F	E	K	=====	D	K	Y	I	L	Q	G	V	T	S	W	G	L	G	=====	C	A	R	P	=====	N	K	P	G	V	Y	V	R	V	=====	S	R	F	V	T	W	I	E	G	V	M	R	N																	
ALP	=====	Y	K	I	C	A	E	H	L	=====	A	R	G	T	D	S	C	Q	D	S	G	G	P	L	V	C	F	E	K	=====	D	K	Y	I	L	Q	G	V	T	S	W	G	L	G	=====	C	A	R	P	=====	N	K	P	G	V	Y	A	R	V	=====	S	R	F	V	T	W	I	E	G	M	M	R	N																								
TPA	=H	L	N	R	T	V	T	D	N	M	L	C	A	G	D	T	R	S	G	P	Q	A	N	L	H	D	A	C	Q	D	S	G	G	P	L	V	C	L	N	D	=====	G	R	M	T	L	V	G	I	S	W	G	L	G	=====	C	G	Q	K	=====	D	V	P	G	V	Y	T	K	V	=====	T	N	Y	L	D	W	I	R	D	N	M	R	P														
UKH	=H	Y	G	S	E	V	T	T	K	M	L	C	A	A	D	P	=====	Q	W	K	T	D	S	C	Q	D	S	G	G	P	L	V	C	S	L	Q	=====	G	R	M	T	L	T	G	I	V	S	W	G	R	G	=====	C	A	L	K	=====	D	K	P	G	V	Y	T	R	V	=====	S	H	F	L	P	W	I	R	S	H	T	K	E	E	N	G	L	A	L											
THR	==S	T	R	I	R	I	T	N	D	M	F	C	A	G	Y	K	P	G	=====	E	G	K	R	D	A	C	E	G	D	S	G	G	P	F	V	M	K	S	P	Y	=====	N	N	R	W	Y	Q	M	G	I	V	S	W	G	E	G	=====	C	D	R	N	=====	G	K	Y	G	F	Y	T	H	V	=====	F	R	L	K	K	W	I	Q	K	V	I	D	R	L	G	S									
CFB	V	K	D	I	S	E	V	T	P	R	F	L	C	T	G	V	S	P	=====	Y	A	D	P	N	T	C	R	G	D	S	G	G	P	L	I	V	H	K	R	=====	S	R	F	I	Q	V	I	S	W	G	V	V	D	V	C	K	N	Q	=====	K	R	Q	K	V	P	A	H	A	R	D	F	I	N	L	F	Q	V	L	P	W	L	K	E	K	L	Q	E	D	L	G	F	L					
CF2	L	T	D	V	R	E	V	T	D	Q	F	L	C	S	G	T	Q	=====	E	D	E	S	P	C	K	E	S	G	A	V	F	L	E	R	R	=====	F	R	F	Q	V	G	L	V	S	W	G	L	N	P	C	L	G	S	A	D	K	N	S	R	K	R	A	P	R	S	K	V	P	P	P	R	O	F	H	I	N	L	F	R	M	Q	P	L	R	Q	H	L	G	D	V	I	N	F	L	P	L
CFD	=====	Y	D	V	L	R	L	M	C	A	E	S	=====	N	R	D	S	C	K	D	S	G	G	P	L	V	C	G	=====	G	V	L	E	G	V	T	S	G	S	R	V	=====	C	G	N	R	=====	K	K	P	G	I	Y	T	R	V	=====	A	T	Y	A	A	W	I	D	H	V	L																													
C1R	G	K	N	R	M	D	V	F	S	O	N	M	F	C	A	G	H	P	=====	S	L	K	Q	D	A	C	Q	D	S	G	G	P	A	V	R	D	P	N	=====	T	D	R	W	A	T	G	I	V	S	W	G	I	G	=====	C	S	R	=====	G	Y	G	F	Y	T	K	V	=====	L	N	Y	V	D	W	I	K	K	E	M	E	E	E																
C1S	A	D	A	E	A	Y	V	T	P	N	M	I	C	A	G	E	=====	K	=G	M	D	S	C	K	D	S	G	G	A	F	A	V	Q	D	P	N	D	K	G	R	F	Y	A	A	G	L	S	W	G	P	Q	=====	C	G	=====	T	Y	G	L	Y	T	R	V	=====	K	N	Y	V	D	W	I	M	K	T	M	Q	E	N	S	T	P	R	E	D													
ASP	=T	Y	H	D	G	V	V	T	I	N	M	M	C	A	E	S	=====	N	R	D	T	C	R	G	D	S	G	S	P	L	V	C	G	=====	D	A	V	E	G	V	T	W	G	S	R	V	=====	C	G	N	G	=====	K	K	P	G	V	Y	T	R	V	=====	S	S	Y	R	M	W	I	E	N	I	T	G	N	M	T	S																			
CAG	=I	F	G	S	Y	D	P	R	R	Q	I	C	V	G	D	R	=====	R	E	R	K	A	A	F	K	G	D	S	G	G	P	L	L	C	N	=====	N	V	A	H	G	I	V	S	Y	G	K	S	=S	G	V	=====	P	P	E	V	F	T	R	V	=====	S	S	F	L	P	W	I	R	T	M	R	S	F	K	L	D	Q	M	E	T	P	L														
TCL	H	Y	N	F	H	P	V	I	G	L	N	M	I	C	A	G	D	L	=====	R	G	K	D	S	C	N	G	D	S	G	S	P	L	L	C	D	=====	G	I	L	R	G	I	T	S	F	G	G	E	K	=C	G	D	R	=====	R	W	P	G	V	Y	T	F	L	S	=====	D	K	H	L	N	W	I	K	I	M	K	G	S	V																	
HAF	H	Y	N	F	N	P	V	I	G	M	N	M	V	C	A	G	S	L	=====	R	G	G	R	D	S	C	N	G	D	S	G	S	P	L	L	C	E	=====	G	V	F	R	G	V	T	S	F	G	L	E	N	K	C	G	D	P	=====	R	G	P	G	V	Y	I	L	L	S	=====	K	K	H	L	N	W	I	M	T	I	K	G	A	V															
CTL	=Y	F	K	N	R	Y	N	K	T	N	Q	I	C	A	G	D	P	=====	K	T	K	R	A	S	F	R	G	D	S	G	G	P	L	V	C	K	=====	K	V	A	A	G	I	V	S	Y	G	Y	K	=D	G	S	=====	P	P	R	A	F	T	K	V	=====	S	S	F	L	S	W	I	K	K	T	M	K	S																						
BTX	=====	A	Y	N	G	L	P	A	K	T	L	C	A	G	V	L	=====	Q	G	G	I	D	T	C	G	D	S	G	G	P	L	I	C	N	=====	G	Q	F	Q	I	L	S	W	G	S	D	P	=C	A	E	P	=====	R	K	P	A	F	Y	T	K	V	=====	F	D	Y	L	P	W	I	Q	I	A	G	N	K	T	A	T	C	P																	
EBP	==V	Y	L	Q	K	V	T	D	V	M	L	C	A	G	E	M	=====	G	G	K	D	T	C	A	G	D	S	G	G	P	L	I	C	D	=====	G	I	L	Q	G	T	S																																																							



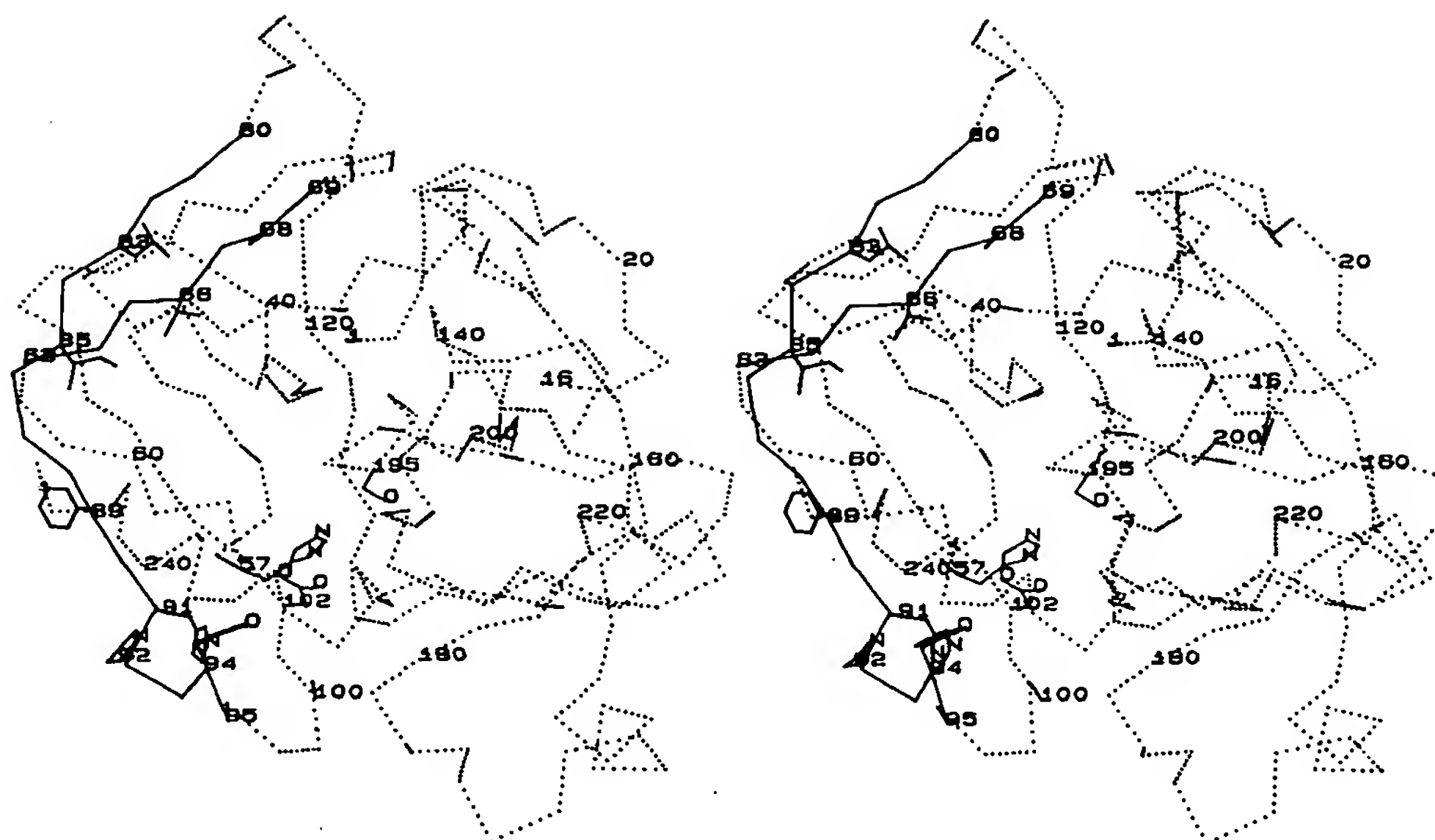


Fig. 5. The location is shown for some of the SCRs that are difficult to align in the serine proteases because of only weak characteristic patterns of homology. This is usually due to the SCR lying on the surface of the molecule. This figure presents the  $C_{\alpha}$  plot of CHT (dotted lines), with the location shown of the respective SCRs that are difficult to align: 63–71 and 81–96 (solid lines). Note that these SCRs lie on the surface of the molecule and

that most of their side chains point out into solvent and thus are free to vary more rapidly than internal residues. In the SCR at 63–71, conserved residues 66 and 68 are the only internal ones and are always hydrophobic; 69 is always a Gly, because it forms a turn. In the SCRs at 81–96, internal hydrophobic residues are 83, 85, 89, and 94. Partial conservation is observed at positions 91 and 92, with a His-Pro.

pletely conserved, probably because it forms a turn and adopts a conformation ( $\phi = \sim 60^\circ$ ,  $\psi = \sim 25^\circ$ ) that is permitted only for Gly (Fig. 5). Similarly, small hydrophobic side chains occur at positions 83 and 85 in the SCR at residues 81–86 (Fig. 5). More varied but always hydrophobic residues appear at 89, and an aromatic one is typical at 94, both of which point into the molecule. In addition, there is also a tendency for a His-Pro at positions 91–92 (Fig. 5); however, a number of sequences have only one or none of these two residues (Fig. 4).

CNO	55	60	65	70	75	80	CNO
	---	>	<--><-->				<-
CHT	AAHCGVT	-----	TSDVVVAGEFDQGSSE	KIQK			CHT
TRP	AAHCYKS	-----	GIQVRLGQDNINV	VEGNQOF			TRP
ELA	AAHCVDR	-----	LTFRVVVGHEHNLNQ	NNGT	EQY		ELA
KAL	AAHCK	-----	NDNYEVLGRHNLFE	NENTAQF			KAL
MCP	AAHCKG	-----	REITVILGAHDVRK	AESTQOK			MCP
SGT	AAHCVSGS	-----	GNNTSITATGGVVDLQS	---	GAAVK		SGT
TON	AAHCYSN	-----	NYOVLLGRNNLEK	DEPFAOR			TON
HPH <sub>old</sub>	TAKNLFLN	=====	HSENATAKDIA	PTLTLY	==	VGKKQL	HPH <sub>old</sub>
HPH <sub>new</sub>	TAKNLFLN	HSENATAKDIA	PTLTLY	VGKK	=====	QL	HPH <sub>new</sub>
	---		<--><-->				<-
CON	AAHC			v	AG		CON

Fig. 6. The old<sup>3,4</sup> and new sequence alignments for HPH, relative to those of the known three-dimensional structures, for the region around the SCR 63–71. Note that six residues have been moved from the VR after position 71 to the VR prior to position 63. Nomenclature and labeling are as in Figure 4.

Fig. 4. Sequence alignment for many of the known serine protease sequences. The source of these various sequences and the definitions of the protein name codes used in this Figure are given in Table II. The first seven proteins (above the line) are those with known experimental three-dimensional structures (see Table I). Their sequences have been aligned based on the superposition of the three-dimensional structures as described in the text. The remaining proteins were aligned using the characteristic sequence matching patterns (see text). When the sequence alignment is uncertain, the respective sequence is shown in italics. The IUPAC-IUB convention standard single letter amino acid code, used in this figure, is as follows: A = Ala, C = Cys, D = Asp, E = Glu, F = Phe, G = Gly, H = His, I = Ile, K = Lys, L = Leu, M = Met, N = Asn, P = Pro, Q = Gln, R = Arg, S = Ser, T = Thr, V = Val, W = Trp, Y = Tyr. An asterisk indicates that the sequence continues past the last residue shown in the figure. Positions of relative deletions in the sequences are denoted by a

dash in the known structures and by a double dash in the sequences aligned using the characteristic homology patterns. The line labeled "CNO" gives the chymotrypsinogen residue numbering used throughout this paper. The  $\leftrightarrow$  symbols delineate the SCRs. The "CON" line lists the conserved, characteristic sequence patterns used to align the "new" sequences (see text). They are coded as follows: upper case, almost completely conserved side chain; lower case, high frequency of this amino acid at this position;  $\lambda$ , nonpolar residues, typically A, V, L, I, M;  $\pi$ , denotes a polar residue, typically S, T, Q, N plus the charged residues; o, denotes S or T, (may be substituted by A occasionally);  $\phi$ , denotes aromatic, usually Y or F, sometimes W; +, denotes a positive residue such as R, K, or H; -, denotes a negative residue such as D or E. The line labeled "S-S" gives the position of the half cystine residue that is disulfide bridged to the half cystine at the label position. A question mark is placed at positions where it is not clear to which residue a disulfide bridge is formed.



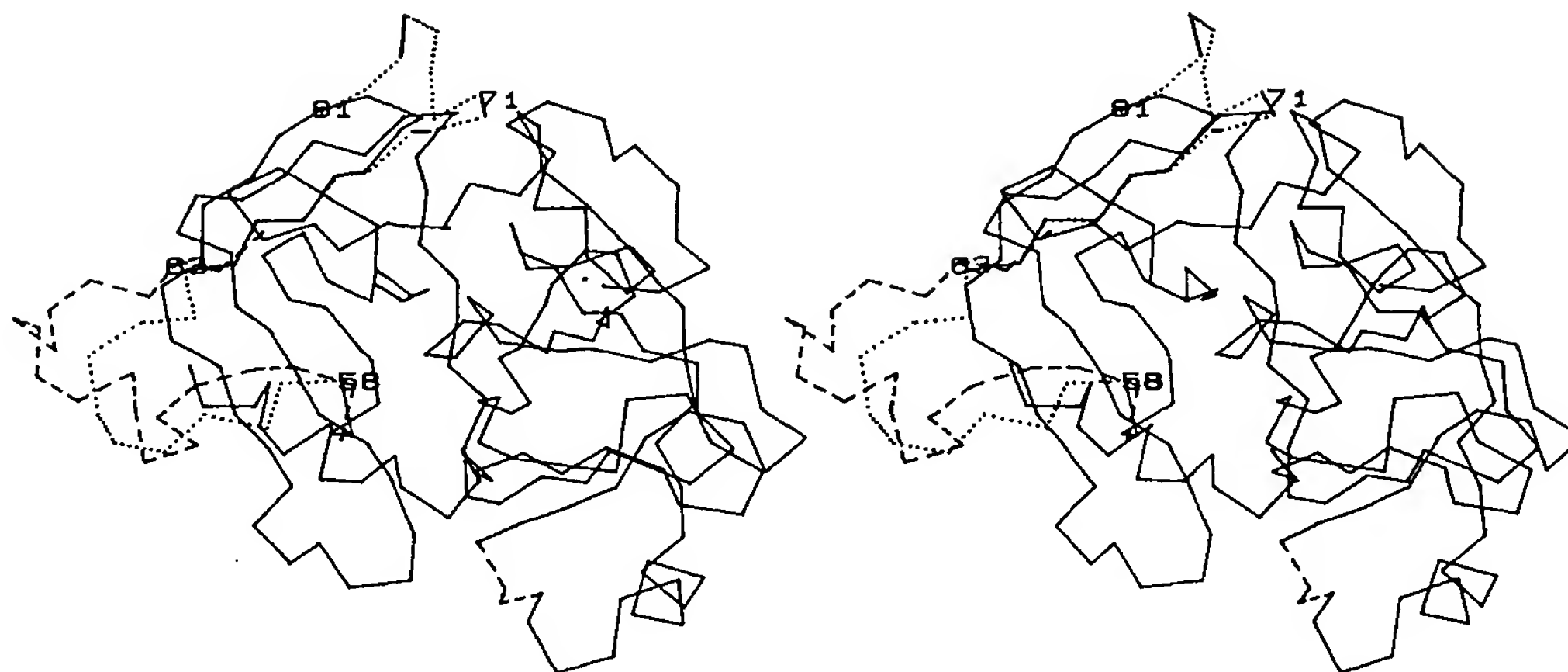


Fig. 7.  $\alpha$  plots for the old (dotted lines) and new (solid lines) structures of HPH based on the respective alignments shown in Figure 6. The differences appear around the SCR at positions 63–71. Because the VR from 59 to 62 in the new HPH alignment is six residues longer than in the previous alignment (Fig. 6) and at least eight residues longer than any of the currently known structures in this loop (see Figs. 4 and 6, Table III), the conformation that has been drawn for this VR (dashed lines) is arbitrary. It has been included only to give a sense of the relatively large size of this loop. The difficulties associated with finding the correct

conformation for such large additions are described in the text. The VR at residues 72–80 is highly truncated, seven residues shorter, in the new structure. Compare the conformations of the known structures for the VRs at 59–62 and 72–80 (Fig. 3) with these loops in HPH. Whereas the  $C\alpha$  positions (and main chain coordinates) for residues 63–71 are the same in this figure for the old and new structures of HPH, the side chains at each residue position are, of course, different in the two alignments (Fig. 6), giving different overall structures for this part of the molecule as well.

### Alignment of “New” Serine Protease Sequences

Identification of the characteristic sequence conservation patterns for each SCR allows the process of aligning “new” sequences to begin. Figure 4 shows the alignment of a large number of serine protease sequences to those of the known structures. The proteins included are taken from a very wide variety of functions and species. It is worth noting that three of these proteins, although clearly homologous to the rest and therefore full members of the family, are no longer functional serine proteases. These are the heavy chain of haptoglobin (HPH), protein Z (PRZ), and the  $\alpha$ -chain of nerve growth factor (NGA). For all these proteins, the characteristic pattern for each SCR was located and matched. Then the remaining positions in that SCR were filled, without permitting any additions or deletions.

One of the results that emerges immediately from this alignment (Fig. 4) is the strong reinforcement of the characteristic patterns in all these “new” sequences. The same basic patterns appear in virtually every protein, yet, in almost every pattern, there are individual protein sequences that have exceptions. Usually the exceptions are minor deviations from the theme of the conservation. For example, the chosen four-residue stretch from 139 to 142, S/T/A-G-W-G, is present in 26 of 35 protein sequences reported in Figure 4. The Trp is replaced by a Phe or a Tyr, both of which are typically accept-

able replacements for a Trp, in four other proteins. In the same way, the characteristic sequence C-G-G almost always appears at positions 42–44, but sometimes one of the Glus becomes an Ala, and occasionally the Cys is replaced by an alternative small side chain. It is clear that, with perhaps the occasional exception of the difficult SCRs between 63–71 and 81–96 discussed above, all the sequences shown can be aligned relatively trivially and unambiguously to the respective characteristic conserved sequences in the SCRs (Fig. 4).

The close examination of the SCRs at 63–71 and 81–96 described in the previous section (Fig. 5), together with the strong reinforcement of the characteristic sequence patterns in these SCRs (Fig. 4), has led to the realignment of the HPH sequence in this region. This results in a large change in the three-dimensional model structure for this protein from that previously proposed.<sup>3,4</sup> The old alignment (Fig. 6) was forced by the need to avoid placing the charged Asp 65 in position 66, where it would be buried in the hydrophobic core of the N-terminal  $\beta$ -barrel. Pro was placed at position 69 as a replacement for a Gly that would permit a turn. Realization that the aliphatic residues at positions 66 and 68 and the almost complete conservation of a Gly at position 69 (Figs. 4 and 5) make up the characteristic sequence pattern for this SCR led to the realignment of the HPH sequence relative to those of the known structures as shown in Figure 6. As a result

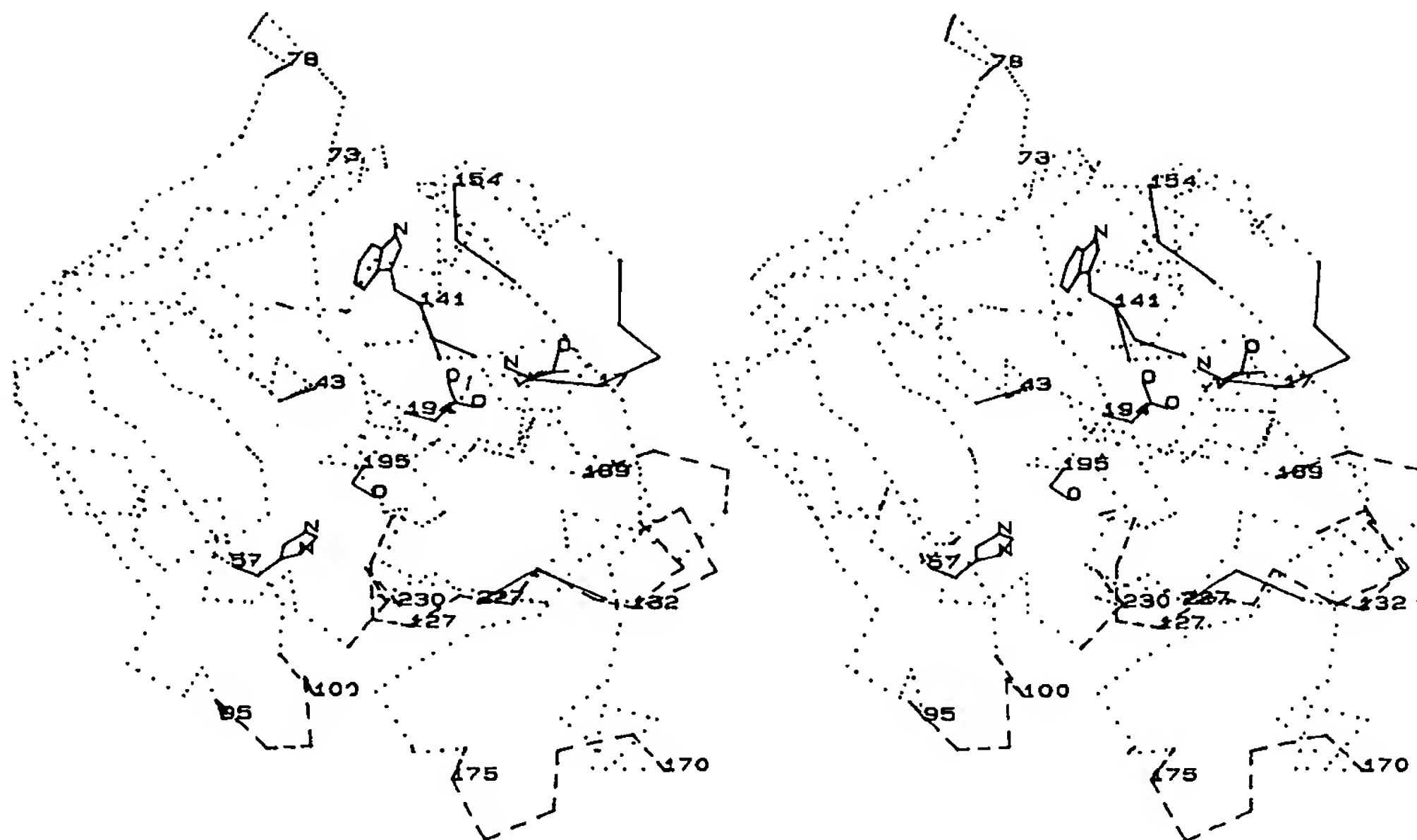


Fig. 8.  $C\alpha$  plot for CHT is shown indicating the sites of VRs with large additions (dashed lines) and with large deletions (dotted lines) in CFB. The unusual residues described in the text are represented by solid lines. These residues include 16–20, 43, 139–142, and 154–156. As can be seen, the unusual residues are localized in one region of the molecule around the normal site

of the N-terminus of the protease peptide chain, residue 16. Residues 16 and 17 should not be in their usual positions, because 16 is not N-terminal in CFB or in CF2. Consequently, they really should be placed in the approximate position expected from our knowledge of the zymogen structures<sup>33–36</sup> where the cleavage at 16 has not yet occurred.

of the new alignment, the HPH sequence has a very large addition in the VR at 59–62 and a very short loop in the VR at 72–80. Now, however, the requirements of aliphatic residues at 66 and 68 and Gly at position 69 are completely satisfied. The difference between the old and new structures for HPH is illustrated in Figure 7, which demonstrates clearly that changes in the alignment can result in large changes to the derived structures.

From an evolutionary perspective, deviation from the typical pattern of sequence conservation occurs most often in those members of the family that are no longer serine proteases in function. For example, HPH and PRZ are no longer serine proteases in function. Both have lost the active site Ser 195 (it is an Ala), and His 57 is also replaced by a Lys in HPH. Nevertheless, both proteins retain most of the characteristic sequence homology patterns, yet they deviate more than the typical sequence. Examples of this are the immediate area around the Ser 195 or the SCR at 63–71. The *Streptomyces griseus* trypsin-like protein also has a weaker homology pattern than the rest, as has been previously noted.<sup>4</sup> Even further diverged are the other bacterial serine proteases.<sup>20–22</sup> These latter are not considered in this paper because of the considerably greater difficulty in predicting their structures correctly by these methods (see footnote \* on page 318).

### Analysis of the Variable Regions

Once the sequences that correspond to the SCRs are identified and aligned, attention can be focussed on the remaining residues forming the VRs. Based on the alignment in Figure 4, the sequence lengths deduced for the various VRs are summarized in Table III. It is clear from Table III that an amazing degree of variation occurs in the length of these loops and therefore in their conformations (Fig. 3).

In previous studies,<sup>4</sup> we classified the modeling of the VRs into five different classes. 1) A known structure has the same-length VR and the same residue character. In this case, a model structure is constructed for the “new” VR by clipping in the respective VR from the known structure. In such cases, the confidence level in this part of the structure can be reasonably good, especially if it is a short VR. One example occurs among the known structures of the serine proteases at residues 97–101, where the VRs of chymotrypsin (CHT), trypsin (TRP), and the mast cell protease (MCP) are all the same length and have the same conformation (see this loop in Fig. 3). 2) Different lengths appear in this VR among the known structures, but they all have the same structural motif, e.g., a  $\beta$ -bend connecting two antiparallel  $\beta$ -strands. The application of a common structural motif to the “new” VR conformation can be illustrated by examining the VR at positions 35–41

in the known structures (Fig. 3). In each case, the known structure consists of two antiparallel  $\beta$ -strands ending in a  $\beta$ -bend, this despite the fact that the loops are very different in length, varying from two to ten residues (Table III). Thus, it would be sensible to model this VR in a "new" structure, at least initially, as a  $\beta$ -bend. Unfortunately, this consideration may not always be sufficient to model the structure accurately enough. For example, we initially modeled this VR in the MCP, which has ten residues, after the conformation of elastase (ELA) in this VR, which also has ten residues.<sup>4</sup> When the crystal structure emerged subsequently,<sup>28</sup> this loop was indeed an antiparallel pair of  $\beta$ -strands ending in a  $\beta$ -bend; however, the strands were bent differently, leading to a quite different conformation for this loop (Fig. 3). This shows that very careful analysis must go into the construction of even this class of loops, especially when the loop to be built is large and has many potential degrees of freedom.

In the next two cases of VRs, the modeling is significantly less reliable. 3) The closest known VR conformation has only a small (i.e., one or two residues) relative addition or deletion to the VR that is being constructed. The addition or deletion of even one residue can lead to considerable changes in the overall conformation of the VR. Examples of this can be seen in Figure 3, where the VR encompassing residues 185–189 is shown. The range of lengths among the known structures is not large for this VR. The shortest is CHT, with three residues; ELA has four; TRP, KAL, MCP, and TON all have five; and SGT has six. As can be seen, the differences between the loops differing by only one residue are considerable, requiring in each case significant rearrangement of the loop conformation. 4) The conformations of the VRs in the known structures vary considerably, even when the sequence length is the same. There are some loops that have different conformations even when the residue lengths are the same. In many cases, it is not easy to identify why the different conformations are occurring. Figure 3 shows an example of such a VR in the 72–80 region of the serine proteases. In cases such as these, construction of the VR is very difficult since we do not know which known structure to use as the example from which to model.

In both the latter cases, a conformational search<sup>41–45</sup> is needed, in principle, to find the correct conformation. Such efforts are very time-consuming and are currently relatively unreliable.<sup>50,51</sup> It is primarily for loops of these two classes that the Protein Data Bank<sup>8</sup> is searched for possible tentative conformations for these loops, as was described above.

The last class of VRs is in some ways the most difficult to construct but perhaps is in other ways easier than previously expected. 5) A very large addition occurs in the VR of the "new" structure. Loops that have large additions are relegated to this class.

Examples of this abound in Figure 4, e.g., a 17 residue addition in the 166–179 loop of HPH or the 11 residue addition in the 97–101 VR of TON relative to CHT. These additions are so large that they preclude the use of even the most current and advanced conformational search methods and energetics analysis. Therefore, in the past,<sup>4</sup> we adopted the practice of not constructing structures for these VRs at all. However, the recently determined experimental structures of KAL<sup>27</sup> and TON<sup>29</sup> show the conformations of two sequences in the VR 97–101 with relatively large additions (Figs. 3 and 4, Table III). In both cases, the loops are disordered in the crystal structures and therefore most probably in solution as well. In the case of KAL, this loop is even cleaved *in vivo* by an exogenous protease in many of the molecules. If this observation of disorder is true, in general, for large additions in the VRs, then it will greatly simplify the modeling of these regions, since there will be no unique structure that has to be determined.

### Sequence Exceptions Warn of Problems

The importance of the characteristic sequence homology to the reliability of the comparative modeling goes beyond the above-described critical task of aligning the sequence of a "new" protein accurately. Specific exceptions to or violations of the characteristic sequences can signal a novel functional property of the protein. Larger discrepancies can be important indicators that the usual extrapolation methods of comparative modeling outlined above are not valid for this particular protein in this specific region.

For this purpose, it is important to distinguish between cases of sequence divergence and violations of the characteristic pattern. Some of the proteins whose sequences are shown in Figure 4, such as HPH, PRZ, and NGA, are no longer serine proteases even though their degree of sequence identity and homology leave no doubt about their being full members of this serine protease structural family. Examination of the characteristic sequence patterns in these proteins shows that they deviate the most from the normal patterns. Nevertheless, clear indication of the pattern usually persists and permits proper, unambiguous alignment of the sequence. For example, the immediate sequence around the active site His 57 and Ser 195 in HPH and PRZ differs more, especially in PRZ, from the usual pattern than any other sequence. However, enough of the pattern remains (even in HPH, where the His and Ser themselves are no longer present) to be clearly recognized and used for sequence alignment.

In other cases, the pattern has faded, and it is difficult to assign the actual alignment. However, the possibilities that are present could probably be accommodated within the rubric of the SCRs of that part of the structure. An example of this appears in



the DSN protein in the SCR from 63 to 71. The usual Gly residue that invariably appears at position 69 is absent. This leads to an ambiguity regarding how to align the sequence. However, at least one of the possibilities, the one shown in Figure 4, is likely to fit into the usual serine protease structure at this position with only minor deviations.

On the other hand, there are examples of larger changes, including the complete absence of one or more characteristic sequence patterns. This can be illustrated by the SCR at 134–145. This SCR has the highly conserved, very characteristic S/T/A-G-W-G sequence at 139–142, as noted above. Virtually every serine protease has this sequence, with a rare change from Trp to Tyr or Phe and a very rare change of the second Gly to something else (see PRZ for an example of divergence as discussed in the previous paragraphs). Note, however, the sequence for this SCR in CFB and CF2 (Fig. 4). In these cases, the characteristic residues are completely missing. Clearly something is happening to the structure here that is unusual. The absence of these residues creates a serious problem in properly aligning these two sequences in this SCR. There is no obvious pattern of residues that would correspond to the S/T/A-G-W-G sequence at 139–142 among the residues that are in this region of the sequence. It is possible, in principle, to place the sequence A-L/H-F-V at these positions, thereby replacing the conserved Trp with a Phe. However, this leaves internal positions 136 and especially 138 with charged residues such as Asp and Lys, where nonpolar residues are characteristic and are called for by the structure. The sequences have been tentatively aligned as shown in Figure 4 to satisfy the requirements that internal residues be nonpolar, or at least uncharged. The uncertain alignments are shown in italics in Figure 4.

If we look further at these two sequences, CFB and CF2, we find another anomalous site. The classic, characteristic pattern for positions 42–44 is C-G-G. Examination of the sequences in Figure 4 shows that substitution of one of the Gly residues by Ala occurs occasionally. The startling introduction of a much larger Met side chain at position 43 in CFB and an even larger and charged Arg in CF2 indicates that once again something strange is happening in these proteins. Examination of the location of these changes in the serine protease three-dimensional structure (Fig. 8) shows that they are immediately adjacent to each other in the structure even though they are quite distant in the sequence.

A further anomalous sequence occurs at positions 16–19 of these two sequences. Activation of a serine protease from its zymogen protein involves the cleavage of the peptide chain prior to residue 16 to generate a free amino terminus at this point in the molecule.<sup>33,52</sup> The earliest serine protease crystal structure solutions<sup>24</sup> showed that this amino terminus formed an internal salt bridge with the side

chain of Asp 194. The homology observed in the residues at positions 16–19 is likely to be diagnostic of this activation step in all these enzymes. Thus it is interesting that several of the sequences in Figure 4 show major deviations from the characteristic sequence pattern for these four residues. All the serine proteases have Ile or Val (or Leu) at positions 16 and 17 and usually Gly at 18 and 19. Exceptions to this pattern are found in PRZ and NGA, neither of which is a true protease and thus presumably no longer retains an activation step. The two proteins CFB and CF2 are also striking in the absence of any homology to the characteristic pattern for this stretch of residues or even between the two themselves. It is known that, unlike the other serine proteases, these two proteins are clipped some 200 residues N-terminal to residue 16.<sup>53,54</sup> Thus the activation mechanism for these proteins must be different from that of the other serine proteases. Figure 8 shows that the normal location of the N-terminal residues 16 and 17 is also close to the two unusual sequences described above for the three-dimensional structure.

Thus significant violations of three spatially adjacent characteristic sequence patterns appear in these two proteins, coupled with functional significance expressed in the absence of the normal mechanism of zymogen activation. It is interesting to note that CFB and CF2 occupy parallel functional roles in the alternative and classical pathways of the complement system, respectively.<sup>55</sup> They form the proteolytic subunits of the respective C3 convertases in the two pathways and, upon binding of C3b subunits, change their specificity to become C5 convertases. Taken together, these observations suggest that some unusual three-dimensional structure is occurring in this particular region of the molecule involving the activation site and the immediately adjacent active and specificity sites. What is difficult to predict or determine is just how different these structures actually are from the typical serine protease structural theme in this region. The excellent preservation of the characteristic sequence homology patterns everywhere else in the CFB and CF2 sequences (Fig. 4) suggests strongly that the remaining parts of the molecule are close to their normal conformation. However, there is no such reassurance for this part of the molecule.

As one begins to construct this part of the CFB molecule onto the typical SCR framework of the serine proteases, several problems are encountered very rapidly. The introduction of side chains at positions 140 and 43 in place of Gly, in such close proximity, causes a steric collision of these two groups. In constructing the N-terminal portion of the chain, the conformation for residues 16–19 must be taken from the zymogen structure of chymotrypsinogen,<sup>33,34</sup> since these residues cannot reside in the usual N-terminal Ile 16 binding pocket. This leaves the problem of what to do with Asp 194, the other



half of the normal salt bridge to the amino terminus at Ile 16. Both possible conformations of the aspartate, that of the zymogen and of the mature enzyme, must be considered based on zymogen crystal studies.<sup>33-36</sup> Unfortunately, neither conformation seems possible without a significant change in the local CFB structure. In the zymogen conformation, the side chains of Met 43 and Ala 140 (both usually glycines) prevent the Asp 194 side chain from occupying its usual place. In the mature enzyme form, it is the side chain of Phe 142 (also typically a glycine) that is too close to permit the usual conformer (see Fig. 8).

The resulting changes that need to be made to accommodate these unusual side chains are difficult to predict. Does the structure undergo a large number of small changes that results in the accommodation of the above groups in close to the normal conformation, or is there a significant change in this portion of the molecule that gives a very different conformation around Asp 194 and residues 139-142? It is not possible to answer this question using the energetics methods currently available. Such problems may have to wait either for experimental structure determinations or for improved understanding of protein structure and energetics. Therefore, it is particularly important that comparative modeling methods, as described in this paper, allow us to recognize such regions of the molecule that are likely to be particularly difficult to construct reliably.

### ACKNOWLEDGMENTS

I am indebted to my colleagues Drs. Stan Burt, John Erickson, and Charles Hutchins for many valuable discussions and for critical reading of the manuscript.

### REFERENCES

- Browne, W.J., North, A.C.T., Phillips, D.C., Brew, K., Vamanam, T.C., Hill, R.L. A possible three-dimensional structure of bovine  $\alpha$ -lactalbumin based on that of hen's egg-white lysozyme. *J. Mol. Biol.* 42:65-86, 1969.
- McLachlan, A.D., Shotton, D.M. Structural similarities between  $\alpha$ -lytic protease of *Myxobacter* 495 and elastase. *Nature [New Biol.]* 229:202-205, 1971.
- Greer, J. Model for haptoglobin heavy chain based upon structural homology. *Proc. Natl. Acad. Sci. USA* 77:3393-3397, 1980.
- Greer, J. Comparative model building of the mammalian serine proteases. *J. Mol. Biol.* 153:1027-1042, 1981.
- Feldmann, R.J., Bing, D.H., Potter, M., Mainhart, C., Furie, B., Furie, B.C., Caporale, L.H. On the construction of computer models of proteins by the extension of crystallographic structures. *Ann. N.Y. Acad. Sci.* 439:12-43, 1985.
- Greer, J. Protein structure and function by comparative model building. *Ann. N.Y. Acad. Sci.* 439:44-63, 1985.
- Blundell, T., Sibanda, B.L., Pearl, L. Three-dimensional structure, specificity and catalytic mechanism of renin. *Nature* 304:273-275, 1983.
- Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Jr., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T., Tasumi, M. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112:535-542, 1977.
- Devereux, J., Haeberli, P., Smithies, O. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12:387-395, 1984.
- Arcoleo, J.P., Greer, J. Hemoglobin binding and its relationship to the serine protease-like active site of haptoglobin. *J. Biol. Chem.* 257:10063-10068, 1982.
- Mollison, K.W., Mandecki, W., Zuiderweg, E.R.P., Fayer, L., Fey, T.A., Krause, R.A., Conway, R.G., Miller, L., Edalji, R.P., Shallcross, M.A., Lane, B., Fox, J.L., Greer, J., Carter, G.W. Identification of receptor-binding residues in the inflammatory complement protein C5a by site-directed mutagenesis. *Proc. Natl. Acad. Sci. USA* 86:292-296, 1989.
- Lustbader, J.W., Arcoleo, J.P., Birken, S., Greer, J. Hemoglobin-binding site on haptoglobin probed by selective proteolysis. *J. Biol. Chem.* 258:1227-1234, 1983.
- Greer, J. Model of a specific interaction: Salt bridges form between prothrombin and its activating enzyme blood clotting factor Xa. *J. Mol. Biol.* 153:1043-1053, 1981.
- Sham, H.L., Bolis, G., Stein, H.H., Fesik, S.W., Marcotte, P.A., Plattner, J.J., Rempel, C.A., Greer, J. Renin inhibitors: design and synthesis of a new class of conformationally restricted analogs of angiotensinogen. *J. Med. Chem.* 31:284-295, 1987.
- Murphy, M.E.P., Moul, J., Bleackley, R.C., Gershenfeld, H., Weissman, I.L., James, M.N.G. Comparative molecular model building of two serine proteinases from cytotoxic T lymphocytes. *Protein Structure Function Genet* 4:190-204, 1988.
- Delbaere, L.T.J., Brayer, G.D., James, M.N.G. Comparison of the predicted model of  $\alpha$ -lytic protease with the X-ray structure. *Nature* 279:165-168, 1979.
- Read, R.J., Brayer, G.D., Jurasek, L., James, M.N.G. Critical evaluation of comparative model building of *Streptomyces griseus* trypsin. *Biochemistry* 23:6570-6575, 1984.
- Zuiderweg, E.R.P., Henkin, J., Mollison, K.W., Carter, G.W., Greer, J. Comparison of model and nuclear magnetic resonance structures for the human inflammatory protein C5a. *Protein Struct Function Genet* 3:139-145, 1988.
- Polgar, L. Serine proteases. In: "Mechanisms of Protease Action." CRC Press, Boca Raton, Florida, 1989: 87-122.
- Fujinaga, M., Delbaere, L.T.J., Brayer, G.D., James, M.N.G. Refined structure of  $\alpha$ -lytic protease at 1.7 Å resolution. Analysis of hydrogen bonding and solvent structure. *J. Mol. Biol.* 184:479-502, 1985.
- Moul, J., Sussman, F., James, M.N.G. Electron density calculations as an extension of protein structure refinement. *Streptomyces griseus* protease at 1.5 Å resolution. *J. Mol. Biol.* 182:555-566, 1985.
- Read, R.J., Fujinaga, M., Sielecki, A.R., James, M.N.G. Structure of the complex of *Streptomyces griseus* protease B and the third domain of the turkey ovomucoid inhibitor at 1.8 Å resolution. *Biochemistry* 22:4420-4433, 1983.
- Read, R.J., James, M.N.G. Refined crystal structure of *Streptomyces griseus* trypsin at 1.7 Å resolution. *J. Mol. Biol.* 200:523-551, 1988.
- Birktoft, J.J., Blow, D.M. Structure of crystalline  $\alpha$ -chymotrypsin. V. The atomic structure of tosyl- $\alpha$ -chymotrypsin at 2 Å resolution. *J. Mol. Biol.* 68:187-240, 1972.
- Chamber, J.L., Stroud, R.M. The accuracy of refined protein structures: comparison of two independently refined models of bovine trypsin. *Acta Crystallogr.* B35:1861-1874, 1979.
- Sawyer, L., Shotton, D.M., Campbell, J.W., Wendell, P.L., Muirhead, H., Watson, H.C., Diamond, R., Ladner, R.C. The atomic structure of crystalline porcine pancreatic elastase at 2.5 Å resolution: Comparison with the structure of  $\alpha$ -chymotrypsin. *J. Mol. Biol.* 118:137-208, 1978.
- Bode, W., Chen, Z., Bartels, K., Kutzbach, C., Schmidt, G., Bartunik, H. Refined 2 Å x-ray crystal structure of porcine pancreatic kallikrein A, a specific trypsin-like serine proteinase. Crystallization, structure determination, crystallographic refinement, structure and its comparison with bovine trypsin. *J. Mol. Biol.* 164:237-282, 1983.
- Remington, S.J., Woodbury, R.G., Reynolds, R.A., Matthews, B.W., Neurath, H. The structure of rat mast cell protease II at 1.9 Å resolution. *Biochemistry* 27:8097-8105, 1988.
- Fujinaga, M., James, M.N.G. Rat submaxillary gland

- serine proteinase, tonin: Structure solution and refinement at 1.8 Å resolution. *J. Mol. Biol.* 195:373–396, 1987.
30. Blevins, R.A., Tulinsky, A. The refinement and the structure of the dimer of  $\alpha$ -chymotrypsin at 1.67 Å resolution. *J. Biol. Chem.* 260:4264–4275, 1985.
  31. Cohen, G.H., Silverton, E.W., Davies, D.R. Refined crystal structure of  $\gamma$ -chymotrypsin at 1.9 Å resolution: comparison with other pancreatic serine proteases. *J. Mol. Biol.* 148:449–479, 1981.
  32. Marquart, M., Walter, J., Deisenhofer, J., Bode, W., Huber, R. The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. *Acta Crystallogr.* B39:480–490, 1983.
  33. Freer, S.T., Kraut, J., Robertus, J.D., Wright, H.T., Xuong, N.H. Chymotrypsinogen: 2.5 Å crystal structure, comparison with  $\alpha$ -chymotrypsin, and implications for zymogen activation. *Biochemistry* 9:1997–2009, 1970.
  34. Wang, D., Bode, W., Huber, R. Bovine chymotrypsinogen A: X-ray crystal structure analysis and refinement of a new crystal form at 1.8 Å resolution. *J. Mol. Biol.* 185:595–624, 1985.
  35. Walter, J., Steigemann, W., Singh, T.P., Bartunik, H., Bode, W., Huber, R. On the disordered activation domain in trypsinogen. Chemical labelling and low temperature crystallography. *Acta Crystallogr.* B38:1462–1472, 1982.
  36. Kossiakoff, A.A., Chambers, J.L., Kay, L.M., Stroud, R.M. Structure of bovine trypsinogen at 1.9 Å resolution. *Biochemistry* 16:654–664, 1977.
  37. Remington, S.J., Matthews, B.W. A systematic approach to the comparison of protein structures. *J. Mol. Biol.* 140:77–99, 1979.
  38. Rossmann, M.G., Argos, P. The taxonomy of protein structure. *J. Mol. Biol.* 109:99–129, 1977.
  39. Waterman, M.S. Computer Analysis of Nucleic Acid Sequences. *Methods Enzymol.* 164:765–793, 1988.
  40. Ponder, J.W., Richards, F.M. Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes. *J. Mol. Biol.* 193:775–91, 1987.
  41. Brucoleri, R.E., Karplus, M. Prediction of the folding of short polypeptide segments by uniform conformational sampling. *Biopolymers* 26:137–168, 1987.
  42. Brucoleri, R.E., Karplus, M. Chain closure with bond angle variations. *Macromolecules* 18:2767–2773, 1987.
  43. Shenkin, P.S., Yarmush, D.L., Fine, R.M., Wang, H., Levinthal, C. Predicting antibody hypervariable loop conformation. I. Ensembles of random conformations for ring-like structures. *Biopolymers* 26:2053–2085, 1987.
  44. Moulton, J., James, M.N.G. An algorithm for determining the conformation of polypeptide segments in proteins by systematic search. *Proteins* 1:146–163, 1986.
  45. Burt, S., Greer, J. Search Strategies for Determining the Bioactive Conformers of Peptides and Small Molecules. *Ann. Rep. Med. Chem.* 23:285–294, 1988.
  46. Hagler, A.T., Stern, P.S., Sharon, R., Becker, J.M. and Naider, F. Computer simulation of the conformational properties of oligopeptides: Comparison of theoretical methods and analysis of experimental results. *J. Am. Chem. Soc.* 101:6842–6852, 1979.
  47. Dauber, P., Osguthorpe, D. and Hagler, A.T. Structure, energetics, and dynamics of ligand binding to dihydrofolate reductase. *Biochem. Soc. Trans.* 10:312–318, 1982.
  48. Kraulis, P.J., Jones, T.A. Determination of three-dimensional protein structures from nuclear magnetic resonance data using fragments of known structures. *Proteins* 2:188–201, 1987.
  49. Struthers, R.S., Hagler, A.T., Rivier, J. In: "Conformationally Directed Drug Design: Peptides and Nucleic Acids as Templates or Targets." Vida, J.A., Gordon, M., eds. Washington, D.C.: Am. Chem. Soc., 1984: 239.
  50. Chothia, C., Lesk, A., Levitt, M., Amit, A., Mariuzza, R., Phillips, V., Poljak, R. The predicted structure of immunoglobulin D1.3 and its comparison with the crystal structure. *Science* 233:755–758, 1986.
  51. Fine, R.M., Wang, H., Shenkin, P.S., Yarmush, D.L., Levinthal, C. Predicting antibody hypervariable loop conformations. II. Minimization and molecular dynamics studies of MCPC603 from many randomly generated loop conformations. *Proteins* 1:342–362, 1986.
  52. Stroud, R.M., Kossiakoff, A.A., Chambers, J.L. Mechanisms of zymogen activation. *Annu. Rev. Biophys. Bioeng.* 6:177–193, 1977.
  53. Mole, J.E., Anderson, J.K., Davison, E.A., Woods, D.E. Complete primary structure for the zymogen of human complement factor B. *J. Biol. Chem.* 259:3407–3412, 1984.
  54. Bentley, D.R., Porter, R.R. Isolation of cDNA clones for human complement component C2. *Proc. Natl. Acad. Sci. USA* 81:1212–1215, 1984.
  55. Reid, K.B.M., Porter, R.R. The Proteolytic activation systems of complement. *Annu. Rev. Biochem.* 50:433–464, 1981.



## Catalytic Domain Structures of MT-SP1/Matriptase, a Matrix-degrading Transmembrane Serine Proteinase\*

Received for publication, October 11, 2001  
Published, JBC Papers in Press, November 5, 2001, DOI 10.1074/jbc.M109830200

Rainer Friedrich‡, Pablo Fuentes-Prior‡, Edgar Ong§, Gary Coombs§, Michael Hunter§,  
Ryan Oehler§, Diane Pierson§, Richard Gonzalez§, Robert Huber‡, Wolfram Bode‡¶,  
and Edwin L. Madison§

From the ‡Max-Planck-Institut für Biochemie, Abteilung Strukturforschung, Am Klopferspitz 18a, 82152 Martinsried,  
Germany and §Corvas International, San Diego, California 92121

The type II transmembrane multidomain serine proteinase MT-SP1/matriptase is highly expressed in many human cancer-derived cell lines and has been implicated in extracellular matrix re-modeling, tumor growth, and metastasis. We have expressed the catalytic domain of MT-SP1 and solved the crystal structures of complexes with benzamidine at 1.3 Å and bovine pancreatic trypsin inhibitor at 2.9 Å. MT-SP1 exhibits a trypsin-like serine proteinase fold, featuring a unique nine-residue 60-insertion loop that influences interactions with protein substrates. The structure discloses a trypsin-like S1 pocket, a small hydrophobic S2 subsite, and an open negatively charged S4 cavity that favors the binding of basic P3/P4 residues. A complementary charge pattern on the surface opposite the active site cleft suggests a distinct docking of the preceding low density lipoprotein receptor class A domain. The benzamidine crystals possess a freely accessible active site and are hence well suited for soaking small molecules, facilitating the improvement of inhibitors. The crystal structure of the MT-SP1 complex with bovine pancreatic trypsin inhibitor serves as a model for hepatocyte growth factor activator inhibitor 1, the physiological inhibitor of MT-SP1, and suggests determinants for the substrate specificity.

The activity of proteolytic enzymes is required at multiple stages during the growth, invasion, and progression of human tumors (for a review, see Ref. 1). For example, these complex processes entail extensive re-modeling of the extracellular matrix as well as the activation of latent growth factors and pro-angiogenic proteins. Consequently, the high level expression of particular proteinases often correlates with poor patient survival for several different cancers (see, for instance, Ref. 2). Among these cancer-associated enzymes are serine proteinases such as urokinase-type plasminogen activator (uPA),<sup>1</sup> elastase,

plasmin, and cathepsin G; matrix metalloproteinases including gelatinases, interstitial collagenases, stromelysins, matrilysin, and membrane-type metalloproteinases; and lysosomal cysteine proteinases such as cathepsin B. Recently, several members of an important, emerging subfamily of serine proteinases, the type II transmembrane serine proteinases (TTSPs; reviewed in Ref. 3), have also been implicated in tumor growth and progression.

MT-SP1 (matriptase/TADG-15/suppressor of tumorigenicity 14; EC 3.4.21) was first isolated by Shi *et al.* (4) as a novel proteinase that was expressed by human breast cancer cells. The enzyme was initially assigned as a gelatinase, because of its gelatinolytic properties and gelatinase-like molecular weight. However, isolation and sequencing of the cDNA revealed a 683-residue multidomain proteinase with a C-terminal serine proteinase domain. The enzyme was then named matriptase to emphasize its matrix degrading properties and trypsin-like specificity (5). Independently, Takeuchi and co-workers (6) cloned and characterized a type-II membrane-bound trypsin-like serine proteinase from a human prostatic cancer cell line, which they called membrane-type serine proteinase 1, MT-SP1. This 855-residue proteinase contained two tandem repeats of the complement component C1r/s domain (CUB, derived from complement factor/1R-urchin embryonic growth factor/bone morphogenetic protein) and four tandem repeats of the low density lipoprotein receptor (LDLR) class A domain between the N-terminal transmembrane signal anchor and the C-terminal catalytic domain (5, 6). Because the matriptase sequence reported by Lin turned out to be part of the translated MT-SP1 cDNA sequence, matriptase is likely to be a form of MT-SP1 produced by ectodomain shedding (7). Alternatively, the two cDNAs may result from alternative splicing. MT-SP1 is highly expressed in prostate, breast, and colorectal cancers *in vitro* and *in vivo* (8), and inhibition of this enzyme suppresses both primary tumor growth and metastasis in a rat model of prostate cancer (5, 6). A mouse homologue was cloned by another group and called epithin (9).

The substrate specificity of MT-SP1 has been mapped using a positional scanning synthetic combinatorial library and substrate phage display (10). The preferred cleavage sequences contained Arg/Lys at P4 and basic residues or Gln at P3, small residues at P2, Arg or Lys at P1, and Ala at P1'. This specificity profile corresponds well to the cleavage sequences of recognized surface localized protein substrates of MT-SP1 such as the proteinase-activated receptor-2 (PAR2), single-chain uPA

\* This work was supported by Grant SFB469 from the University of Munich, by Grant ERBFMRXCT98 from the Training and Mobility Program of the European Union, and by the Fonds der Chemischen Industrie. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1EAX (Bz-MT-SP1) and 1EAW (BPTI-MT-SP1)) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

¶ To whom correspondence should be addressed. Tel.: 49-89-8578-2676; Fax: 49-89-8578-3516; E-mail: bode@biochem.mpg.de.

<sup>1</sup> The abbreviations used are: uPA, urokinase-type plasminogen activator; sc-uPA, single-chain urokinase-type plasminogen activator;

TTSP, type II transmembrane serine proteinase; LDLR, low density lipoprotein receptor; HAI, hepatocyte growth factor activator inhibitor; PAR, proteinase-activated receptor; Bz, benzamidine; BPTI, bovine pancreatic trypsin inhibitor.

(sc-uPA), the proform of MT-SP1, and the hepatocyte growth (scattering) factor, which have been shown *in vitro* and/or *in vivo* to be efficiently activated by MT-SP1 (10, 11).

Although human breast cancer cells produce MT-SP1 primarily as the free enzyme, in human milk and normal tissues the enzyme is found in complex with an inhibitor called hepatocyte growth factor activator inhibitor 1 (HAI-1; Ref. 12). This membrane-bound inhibitor was originally isolated from human stomach carcinoma cells (13) as a 478 residue glycoprotein containing two Kunitz-type domains separated by an LDLR domain and followed by a transmembrane segment, but has been subsequently detected in several tissues. Soluble, presumably proteolytically cleaved forms of HAI-1, lacking the C-terminal hydrophobic domain, have also been reported (14). In addition to HAI-1, a smaller inhibitor (HAI-2) has been identified and characterized, which lacks the LDLR domain separating the two Kunitz-type modules in HAI-1. Site-directed mutagenesis studies suggested that the first Kunitz domain of HAI-2 is responsible for the inhibitory activity toward hepatocyte growth factor activator (15).

We have expressed and purified the catalytic domain of human MT-SP1 (MT-SP1(cd)), and we have crystallized and solved the high resolution x-ray crystal structure of this enzyme in the presence of benzamidine (Bz). Because recombinant HAI-1 was not available, we also determined the structure of the MT-SP1 complex with the Kunitz-type bovine pancreatic trypsin inhibitor (BPTI), which shares a 36% sequence identity with the first Kunitz domain of HAI-1 and is a nanomolar range inhibitor of MT-SP1. These crystal structures provide important new insights into the molecular determinants of the unique specificity of MT-SP1 not obtainable from modeling (16), and give hints about the interaction of this proteinase with the physiological inhibitor HAI-1. This information is expected to facilitate the design of potent, selective small molecule inhibitors of MT-SP1 that may yield lead compounds for the development of novel anti-cancer agents. Because of the high accessibility of the substrate binding site, the MT-SP1(cd) crystals are well suited for soaking of small molecule inhibitors facilitating the further elaboration of initial lead compounds.

#### EXPERIMENTAL PROCEDURES

**Cloning and Purification**—The human prostate adenocarcinoma cell line, PC-3, was purchased from ATCC (CRL-1435). PC-3 cells were lysed in Trizol reagent (Invitrogen, Carlsbad, CA) and total RNA was isolated according to the manufacturer's protocol. Poly(A)<sup>+</sup> RNAs were purified using oligo(dT) beads (Oligotex; Qiagen, Valencia, CA) and subsequently converted to single-stranded cDNAs by reverse transcription using ProSTAR first-strand reverse transcriptase-PCR kit (Stratagene, La Jolla, CA) and SuperScript II RNase H<sup>-</sup> reverse transcriptase (Invitrogen). Single-stranded cDNAs from PC-3 cell RNA were subjected to PCR with sense and antisense degenerate oligonucleotide primers (sense primer, 5'-TGGRT(I)VT(I)WS(I)GC(I)RC(I)CAYTG-3'; antisense primer, 5'-(I)GG(I)CC(I)CC(I)SWRTC(I)CCYT(I)RCA(I)GHRTC-3', where R = A or G; V = G, A, or C; W = A or T; S = G or C; Y = C or T; and H = A, T, or C). The primer sequences corresponded to two highly conserved regions in all chymotrypsin-like serine proteinases. PCR products were purified using a gel extraction kit (QIAquick gel extraction kit; Qiagen), ligated into pCR2.1-TOPO (Invitrogen), and transformed into *Escherichia coli* TOP10 cells (Invitrogen). To obtain additional MT-SP1 cDNA sequences, both rapid amplification of cDNA ends and gene-specific amplification reactions were performed. A human prostate Marathon-Ready cDNA (CLONTECH, Palo Alto, CA) was used to isolate part of the cDNA encoding MT-SP1. The 3' region of MT-SP1 cDNA was successfully obtained by a 3'-rapid amplification of cDNA ends reaction using a gene-specific primer, 5'-CACCCCTTCTTCAATGACTTCACCTTCG-3'. The 5' end of the MT-SP1 proteinase domain was obtained by a PCR amplification reaction using two MT-SP1-specific primers, 5'-TACCTCTCTACGACTCC-3' for the sense primer and 5'-GAGGTTCTCGCAGGTGGTCTGGTTG-3' for the antisense primer. These fragments were subcloned into pCR2.1-TOPO.

After transformation into *E. coli* cells, the insert DNAs were characterized by Southern blot analysis (using the internal cDNA fragment as probe) and by DNA sequence analysis. To obtain a cDNA encoding the entire proteinase domain of MT-SP1, an end-to-end PCR amplification using the gene-specific primers 5'-TCTCTCGAGAAAAGAGTTGTTGGGGGCACGGATGCGGATGAG-3' for the 5' end and 5'-ATTCCGCGCCGCCTATACCCAGTGTCTCTTTGATCCA-3' for the 3' end. An 800-bp DNA fragment was amplified, purified, digested with *Xho*I and *Not*I, and subcloned into the *Pichia pastoris* expression vector, pPIC9KX. Transformation was performed in a Bio-Rad GenePulser II (voltage = 1500 V, capacity = 50 microfarads, and resistance = 200 ohms). The screening of transformed *Pichia* clones for MT-SP1 expression was performed by testing clones with Spectrozyme t-PA (CH<sub>3</sub>SO<sub>3</sub>-D-HHT-Gly-Arg-pNA.HCl; American Diagnostica).

The production of multimilligram amounts of MT-SP1 was carried out by fermentation in a BioFlo 3000 fermentor (New Brunswick Scientific, NJ) using a SMD1168/pPIC9K:MT-SP1 Sac SC1 clone. The medium was inoculated with an overnight culture of the *P. pastoris* transformant. Cells and cell debris were removed by centrifugation, the supernatant was concentrated, and the buffer was exchanged into 50 mM Tris-HCl, 50 mM NaCl, 0.05% Tween 80, pH 8.0 (buffer A). The concentrated MT-SP1-containing solution was applied onto a 150-ml benzamidine column equilibrated with buffer A, and the column was washed with 50 mM Tris-HCl, 1.0 M NaCl, 0.05% Tween 80, pH 8.0 (buffer B), and eluted with 50 mM Tris-HCl, 1.0 M L-arginine, 0.05% Tween 80, pH 8.0 (buffer C). Fractions containing MT-SP1 activity were pooled and concentrated. The buffer was exchanged into 50 mM Na<sub>2</sub>HPO<sub>4</sub>, 125 mM NaCl, pH 5.5 (buffer D), and the partially purified MT-SP1 was passed through a Q-Sepharose Fast Flow HiTrap column (Amersham Biosciences, Inc.) pre-equilibrated with buffer D. The flow-through was collected, and the protein concentration was determined by measurement of A<sub>280</sub> (using an extinction coefficient of 2.012 mg/A<sub>280</sub>). Purified MT-SP1 was deglycosylated with endoglycosidase H (ProZyme, 5 units/ml) and further purified on a Äkta Explorer system using a 7-ml Source 15Q anion exchange column (Amersham Biosciences, Inc.). The protein was eluted in a buffer containing 50 mM HEPES, pH 6.5, with a 0–0.33 M NaCl gradient. Fractions containing protein were pooled, and benzamidine was added to a final concentration of 10 mM. The protein purity was examined by SDS-PAGE, and the protein concentration was determined by measurement of A<sub>280</sub>.

**Crystallization, Structure Determination, and Crystallographic Refinement**—Plate-like crystals of the Bz-MT-SP1 complex were grown from 0.1 M Tris-HCl, pH 8.0, 1.5 M ammonium sulfate, 3% ethanol at 18 °C using the hanging drop vapor diffusion technique. These crystals belong to the orthorhombic space group C<sub>222</sub>, diffract x-rays to beyond 1.3-Å resolution, and have one molecule in the asymmetric unit. The Bz-MT-SP1 crystals were transferred to 0.1 M Tris-HCl, pH 8.0, 1.5 M ammonium sulfate, 23% glycerol. A complete native data set to 1.3-Å resolution was collected from a single crystal under a nitrogen stream at 100 K using synchrotron radiation and a CCD system (MAR Research, Hamburg, Germany) at DESY, Hamburg, Germany.

These data were evaluated with the MOSFLM package (43) and loaded and scaled using SCALA from the CCP4 program suite (17). For the determination of the orientation and position of the MT-SP1 molecules in the crystals, rotational and translational searches were performed with AMoRe (18) using data from 20- to 3.5-Å resolution and a modified enteropeptidase search model (19) with all nonidentical residues reduced to Ala. A unique solution was found with a correlation factor of 39.4% and an *R*-factor of 46.2%; the corresponding values of the next best solution were 12.8 and 56.3%, respectively. Crystallographic refinement was done in several cycles consisting of model building performed with MAIN (20) and conjugate gradient minimization and simulated annealing using CNS (21). The target parameters of Engh and Huber (22) were used. This procedure converged rapidly, yielding a model with excellent parameters (see Table I). In the final model building/refinement cycles, water molecules were inserted at stereochemically reasonable sites, and individual restrained atomic *B*-values were refined. 5.1% of all reflections were omitted from the refinement to calculate the *R*<sub>free</sub>; the final *R* and *R*<sub>free</sub> are 18.4 and 19.3%, respectively, for all data to 1.3 Å. The whole main chain of the MT-SP1 catalytic domain is in appropriate electron density. Only a few side chains projecting out into solution are partially undefined in the electron density; the occupancy of all undefined atoms was set to zero. Modeling was performed interactively using MAIN; these models were energy refined with CNS.

Crystals of the complex were grown from 0.1 M Hepes, pH 6.5, 20% polyethylene glycol 4000, 2% CsCl<sub>2</sub> at 18 °C using the hanging drop vapor diffusion technique. These crystals belong to the triclinic space



TABLE I  
Crystallographic data

Inhibitor	Bz	BPTI
Data collection		
Space group	C222	P1
Molecules in asymmetric unit	1	2
Cell constants		
<i>a</i>	66.92 Å	47.10 Å
<i>b</i>	141.60 Å	54.23 Å
<i>c</i>	51.94 Å	67.82 Å
$\alpha$	90.00°	107.62°
$\beta$	90.00°	96.86°
$\gamma$	90.00°	103.36°
Limiting resolution	1.30 Å	2.93 Å
<i>R</i> <sub>merge</sub> , overall	8.6%	12.0%
<i>R</i> <sub>merge</sub> , outermost shell	29.4%	34.3%
Unique reflections	58,805	12,200
Completeness, overall	96.2%	94.4%
Completeness, outermost shell	93.5%	93.1%
Non-hydrogen protein atoms	1864	4618
Heterogen atoms	14	0
Solvent atoms	381	80
Refinement statistics		
Test set size	5.1%	8.1%
Resolution range	18.0–1.30 Å	12.0–2.93 Å
Completeness, overall	96.2%	94.4%
Completeness, outermost shell	91.7%	93.1%
<i>R</i>	18.4%	19.9%
<i>R</i> , outermost shell	23.7%	30.0%
<i>R</i> <sub>free</sub>	19.3%	27.9%
<i>R</i> <sub>free</sub> , outermost shell	25.0%	40.8%
Root mean square standard deviations		
Bond length	0.010 Å	0.007 Å
Bond angles	1.6°	1.4°
Average <i>B</i> -value/S.D.	13.7 Å <sup>2</sup>	14.4 Å <sup>2</sup>
Ramachandran plot		
Most favored region	89.3%	74.3%
Favored region	10.7%	23.5%
Generously allowed region		2.3%

group P<sub>1</sub>, diffract x-rays to beyond 2.9-Å resolution, and have two molecules in the asymmetric unit. After transfer of one crystal to its mother solution containing 20% glycerol, a complete native data set to 2.93-Å resolution was collected from a single crystal under a nitrogen stream at 100 K using an Image Plate system (MAR Research). The data were processed as described above, with a combined model of the BPTI-trypsin complex (23) and our coordinates as replacement input. The procedure converged rapidly, leading to *R* and *R*<sub>free</sub> of 19.9 and 26.9%, respectively. 8.1% of the reflections were omitted from the refinement for the calculation of the *R*<sub>free</sub>.

## RESULTS

**Overall Structure of the MT-SP1 Catalytic Domain**—Our cloning, expression, and purification procedure described under "Experimental Procedures" yielded multimilligram quantities of highly purified MT-SP1(cd), which in the presence of Bz formed crystals diffracting to beyond 1.3-Å resolution. MT-SP1(cd) resembles an oblate ellipsoid with diameters of 35 and 50 Å. Similar to other trypsin-like serine proteinases, the chain is folded into two adjacent six-stranded  $\beta$ -barrels strapped together by three trans-domain segments. The surface contains several turn structures, a  $3_{10}$ -helix (residues 60(I) to 64, using the chymotrypsinogen numbering, see Figs. 1A and 2), and two  $\alpha$ -helices (segments 164–172 and 235–243). The catalytic triad is located along the junction of the barrels, whereas the active site cleft runs perpendicular to this junction.

Recombinant MT-SP1(cd) consists of the B-chain of mature MT-SP1. The chain starts with Val<sup>16</sup> (corresponding to Val<sup>617</sup>(g) in the generic MT-SP1 sequence numbering; Ref. 6). Cys<sup>122</sup>, which would be disulfide-linked with Cys<sup>1</sup> of the A-chain in the full-length MT-SP1 molecule, is an unpaired surface located cysteine residue in this construct. The MT-SP1 B-chain contains three disulfide bridges (Cys<sup>42</sup>-Cys<sup>58</sup>, Cys<sup>168</sup>-

Cys<sup>182</sup>, and Cys<sup>191</sup>-Cys<sup>220</sup>) that are also present in most other trypsin-like serine proteinases. However, based on the electron density of the 1.3-Å Bz-MT-SP1(cd) structure, the disulfide bridge Cys<sup>42</sup>-Cys<sup>58</sup> is present in only about half of the crystalized protein molecules. In the other half, these cysteines clearly exist in the reduced form, with Cys<sup>58</sup> S $\gamma$  located in the same position as in the Cys<sup>42</sup>-Cys<sup>58</sup> disulfide, and Cys<sup>42</sup> S $\gamma$  rotated toward the interior of the molecule between Tyr<sup>59</sup> O $\gamma$  and the Leu<sup>33</sup> side chain, thereby avoiding steric hindrance with the adjacent Cys<sup>58</sup> thiol group. A similar partial opening of the Cys<sup>42</sup>-Cys<sup>58</sup> disulfide bridge has recently been found in a high resolution structure of the recombinant human uPA catalytic domain (24).

The  $\alpha$ -ammonium group of the N-terminal Val<sup>16</sup> of MT-SP1 forms the highly conserved internal salt bridge with the side chain carboxylate of Asp<sup>194</sup>, stabilizing the substrate binding site and the active site in the catalytically active conformation. In contrast to most (chymo)trypsin-like proteinases, the whole C-terminal region of MT-SP1, including the last residue, Val<sup>244</sup>, is fully defined by electron density. After the conserved C-terminal  $\alpha$ -helix, the MT-SP1 polypeptide makes a sharp turn at Gly<sup>243</sup> and forms an as yet unobserved surface-located salt bridge with the Arg<sup>235</sup> guanidyl group via its C-terminal carboxylate group.

An optimal superposition with several related serine proteinases reveals highest topological similarity of MT-SP1(cd) with the catalytic domain of another membrane-type serine proteinase, enteropeptidase/enterokinase (Fig. 1B). 222 C $\alpha$  atoms of topologically equivalent residues are found within a 2.0-Å distance, corresponding to an root mean square deviation of 0.70 Å, with 109 of these topologically equivalent residues being identical. The next best fit, with a 0.73-Å root mean square deviation for 212 C $\alpha$  atoms, is observed with bovine trypsin (25), followed by bovine chymotrypsin (26) and human thrombin (27). The topological equivalence with chymotrypsin(ogen) formed the basis for the sequence alignment and the chymotrypsinogen numbering of the MT-SP1 catalytic domain used in this paper (Fig. 2). This alignment requires a nine-residue insertion between residues Ile<sup>60</sup> and Pro<sup>61</sup>, single-residue insertions behind residues Gly<sup>184</sup>, Glu<sup>186</sup>, Ala<sup>204</sup>, and Ala<sup>221</sup>, and single-residue deletions at positions 149 and 218. Inserted residues are marked by suffixes following the residue number of the preceding common residue.

**Loops Surrounding the Active Site Cleft**—The narrow substrate specificity of MT-SP1 arises from unique structural determinants of its active site cleft that is shaped in part by the surrounding surface loops. In the following brief description, these loops (defined according to the residue number of their central residue) will be addressed in an anticlockwise manner with respect to the standard orientation displayed in Fig. 1A.

To the east of the catalytic triad, the rigid 37 loop projects out of the molecular surface of MT-SP1. This loop contains two of the "zymogen triad" residues, Ser<sup>32</sup> and His<sup>40</sup>, which, together with Asp<sup>194</sup>, would stabilize the inactive zymogen-like conformation of the proenzyme (28, 29). Around Gln<sup>38</sup>, this loop deviates markedly from the path followed in most other chymotrypsin-like proteinases (Fig. 1B). This conformation seems to be mainly stabilized by the Gln<sup>38</sup> side chain, which is held in an exposed position via hydrogen bonds made through its terminal carboxamide group with Tyr<sup>60</sup>(G) O $\gamma$  and with the  $\pi$ -electron system of the Phe<sup>60</sup>(E) phenyl ring.

The most striking feature of MT-SP1 is the unusually large 60 insertion loop, which is of the same length (nine additional residues) and exhibits a similar  $\beta$ -hairpin conformation as the corresponding loop in thrombin (27), but is oriented differently (Fig. 1B). The eight residues of MT-SP1 between Tyr<sup>59</sup> and

FIG. 1. Ribbon stereo plots of MT-SP1(cd) shown in standard orientation. Figure was prepared with SETOR (38). A, the fold, secondary structure elements, and loops of MT-SP1. Helices and  $\beta$ -strands are given as red helices and blue strands. The active site residues, Asp<sup>189</sup>, and the disulfide bridges (yellow) are shown as stick models. B, MT-SP1 (red) superimposed with the enteropeptidase catalytic domain (green; Ref. 19), bovine chymotrypsin A (yellow; Ref. 39), and the heavy chain of human  $\alpha$ -thrombin (blue; Ref. 27). The active site residues of MT-SP1, Asp<sup>189</sup>, and the bound benzamidine are shown as stick models.

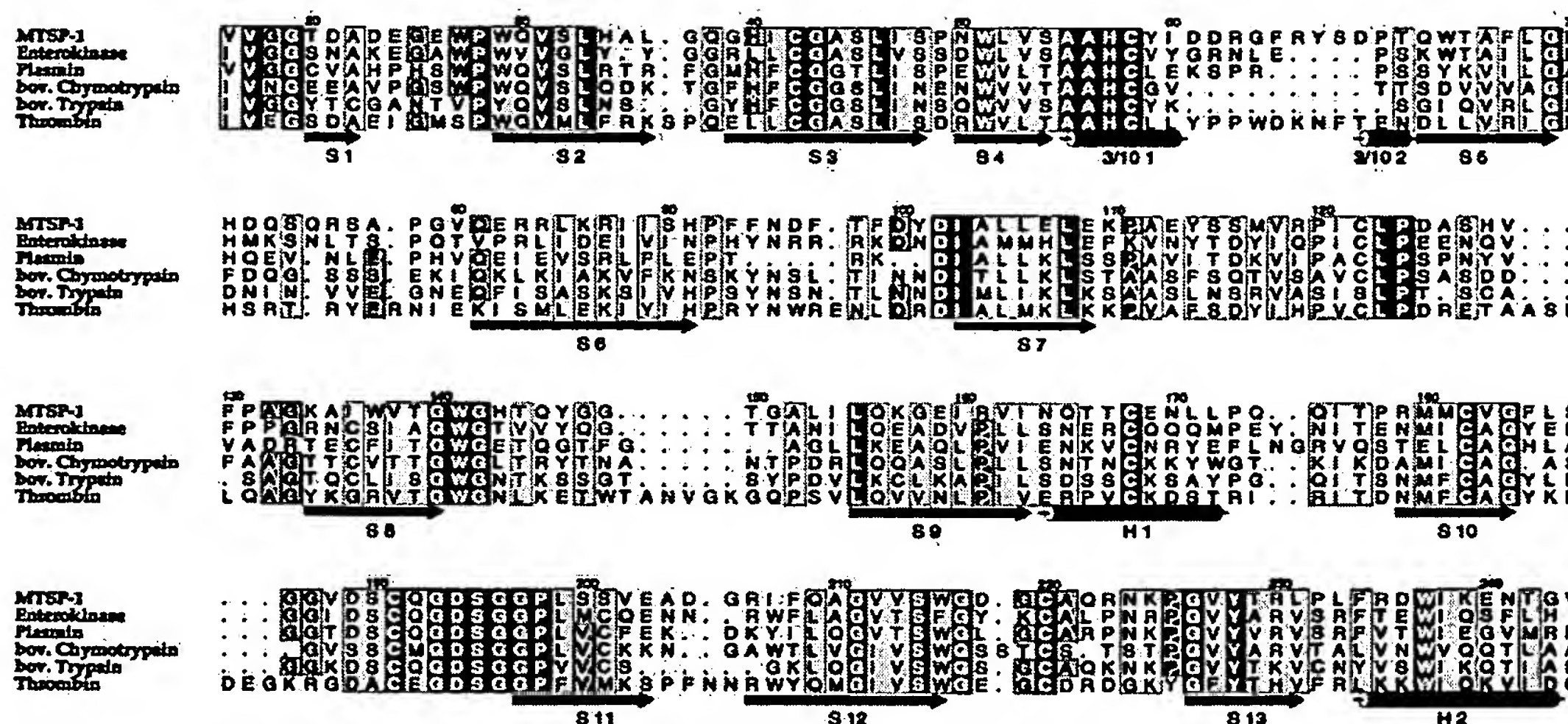
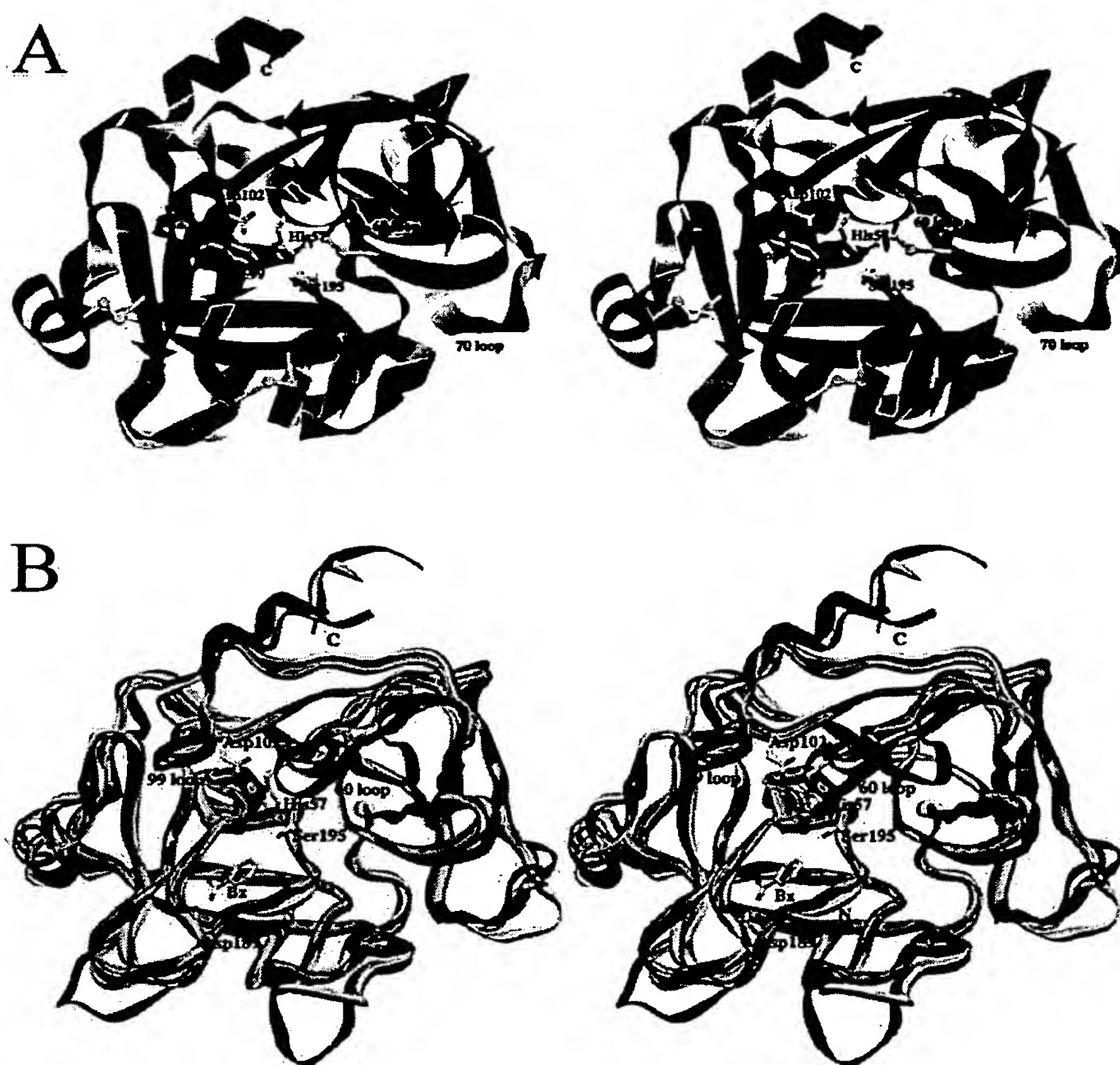


FIG. 2. Structure-based amino acid sequence alignment of the human MT-SP1 catalytic domain with human enteropeptidase (40), human plasmin, bovine chymotrypsinogen A, bovine trypsin, and human thrombin. The numbers given correspond to the chymotrypsinogen numbering.  $\beta$ -Sheets and  $\alpha$ -helices in MT-SP1 are indicated by arrows and cylinders, respectively. Figure was made with ALSCRIPT (41).

Ser<sup>60</sup>(H) form a protruding irregular  $\beta$ -hairpin loop, which is stabilized through internal main chain-side chain hydrogen bonds made by the Asp<sup>60</sup>(A) and Asp<sup>60</sup>(B) carboxylate groups. In thrombin, the rigid 60 loop partially occludes the active site and hence contributes to the narrow substrate specificity (27). In MT-SP1, the loop is rotated away from the active site creating a more open cavity than in thrombin. Unlike the basic side chains of Lys<sup>60</sup>(F) in thrombin or Arg<sup>60</sup>(G) in enteropepti-

dase (19), the Arg<sup>60</sup>(F) side chain of MT-SP1 is directed away from the active site cleft and located between the side chains of Asp<sup>60</sup>(A) and Asp<sup>60</sup>(I).

To the north of the active site, the 99 loop of MT-SP1 protrudes from the molecular surface and forms a "roof" on top of the active site cleft. Noteworthy is the clustering of aromatic side chains around this loop. The benzyl side chain of Phe<sup>99</sup>, which protects the hydrogen bonding interaction between the



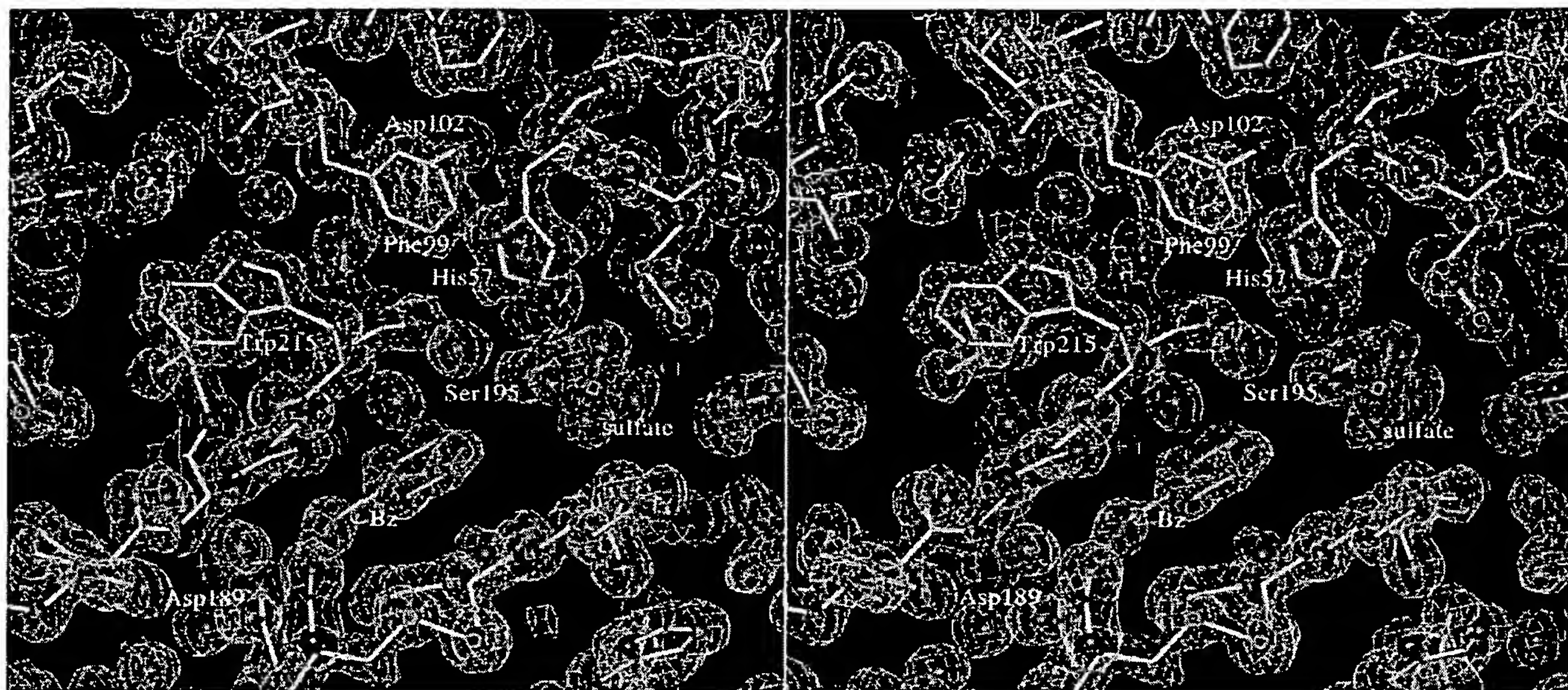


FIG. 3. Stereo section of the Bz-MT-SP1 structure around the S1 pocket, superimposed with the final 1.3-Å  $2F_{\text{obs}} - F_{\text{calc}}$  electron density map contoured at  $1.0\sigma$ . The MT-SP1 model is slightly rotated compared with the standard orientation shown in Fig. 1. Figure was made with MAIN (20).

catalytic residues Asp<sup>102</sup> and His<sup>57</sup> from the solvent, delimits the S2 subsite. Unique to MT-SP1 is the well defined benzyl side chain of Phe<sup>97</sup>, which extends away from the apex of the 99 loop partially shielding the circularly arranged carbonyl groups of Asp<sup>96</sup>, Phe<sup>97</sup>, and Thr<sup>98</sup> (see Fig. 3).

The "autolysis" or 145 loop of MT-SP1 is one residue shorter and less exposed than the corresponding segment of chymotrypsin. This loop encircles the extended side chain of His<sup>143</sup> and forms the southern boundary or "floor" of the active site cleft (Fig. 4). The presence of Gly<sup>151</sup> in this loop is noteworthy, because this would allow for the accommodation of bulkier P2' side chains of peptide substrates.

In the pancreatic serine proteinases, the 70–80 loop forms the calcium binding site, with the carboxylates of Glu<sup>70</sup> and Glu<sup>80</sup> coordinating the calcium ion. In MT-SP1, both positions are occupied by aliphatic residues (Leu<sup>70</sup> and Val<sup>80</sup>), hence rendering the catalytic domain calcium-independent. Together with the proximal side chain parts of Arg<sup>76</sup> and Ala<sup>77</sup>(A), the hydrophobic side chains at positions 70 and 80 clamp both loop ends together in a manner similar to that for the ion metal in the calcium-containing serine proteinases.

**Active Site and Substrate Binding Sites of MT-SP1**—The residues of the active site triad, Ser<sup>195</sup>, His<sup>57</sup>, Asp<sup>102</sup>, and other catalytic elements such as the oxyanion hole created by the main chain nitrogens of Gly<sup>193</sup> and Ser<sup>195</sup> are arranged in the active site cleft exactly as in trypsin and chymotrypsin. The specificity pocket S1, which opens to the west of Ser<sup>195</sup>, is bordered by segments Val<sup>213</sup>–Gly<sup>220</sup>, Ser<sup>190</sup>–Ser<sup>195</sup>, Pro<sup>225</sup>–Tyr<sup>228</sup>, and the Cys<sup>191</sup>–Cys<sup>220</sup> disulfide bridge. The side chains of all MT-SP1 residues lining the interior of this pocket with their side chains are virtually superimposable with the corresponding residues in trypsin. This applies not only to Asp<sup>189</sup> at the bottom of the pocket, determining the specificity for basic residues, but also to Ser<sup>190</sup>, making Lys P1 residues equally acceptable as Arg, and residues forming the inner wall (Gly<sup>226</sup>, Tyr<sup>228</sup>, and Val<sup>213</sup>). The phenyl group of the bound benzamidine molecule is sandwiched between the parallel peptide groups Trp<sup>215</sup>–Gly<sup>216</sup> and Cys<sup>191</sup>–Gln<sup>192</sup>, whereas its amidino group opposes the Asp<sup>189</sup> carboxylate at the bottom of the pocket forming a two-O/two-N salt bridge, with the distal ni-

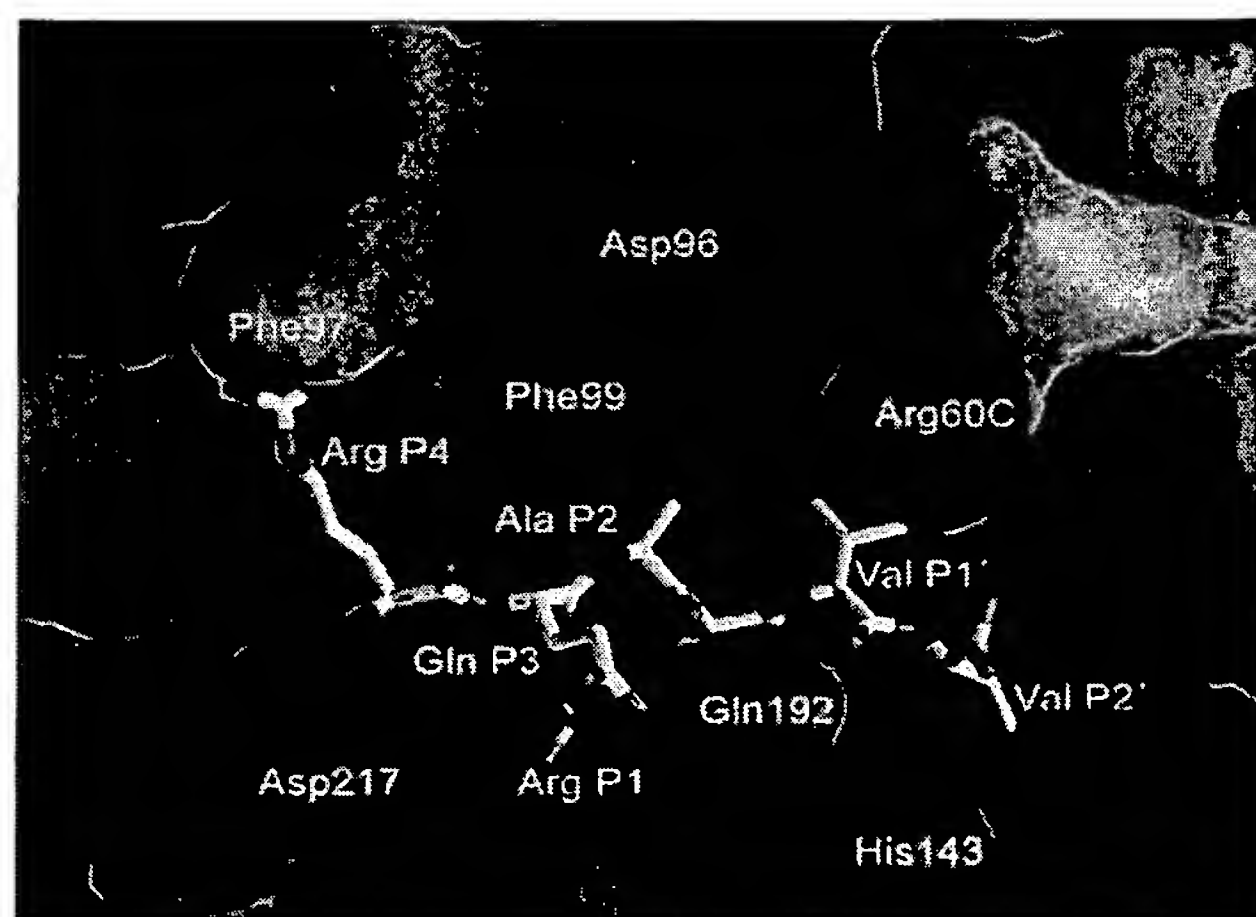


FIG. 4. An Arg-Gln-Ala-Arg↓Val-Val hexapeptide substrate with the sequence of the autoactivation site of pro-MT-SP1, modeled in canonical conformation against the MT-SP1 active site, based on protein inhibitor-trypsin complexes (30). MT-SP1 is represented by its solid surface, with the colors indicating positive and negative electrostatic potential contoured from 15 kT/e (intense blue) to –15 kT/e (intense red). Figure was made with GRASP (42).

trogen atom additionally hydrogen-bonded to Gly<sup>219</sup> O and to a buried solvent molecule.

An extended S2/S4 pocket based on the indole moiety of Trp<sup>215</sup> is located above the S1 pocket. This pocket is delimited to the east by the flat side of the His<sup>57</sup> imidazole ring and to the north by the flat side of the Phe<sup>99</sup> benzyl moiety, and shaped to accept small hydrophobic P2 residues (Fig. 4). The adjacent S4 pocket is bordered by the carbonyl groups of Asp<sup>96</sup>, Phe<sup>97</sup>, and Thr<sup>98</sup>, which are partially shielded from solvent by the benzyl group of Phe<sup>97</sup>, a residue rarely present at this position in (chymo)trypsin-like proteinases. Because of the presence of the  $\pi$ -electron system of Phe<sup>99</sup> and the nearby Asp<sup>96</sup>, this S4 pocket is well suited to accommodate positively charged side chains. It is noteworthy that the benzyl group of Phe<sup>97</sup> is fully defined in



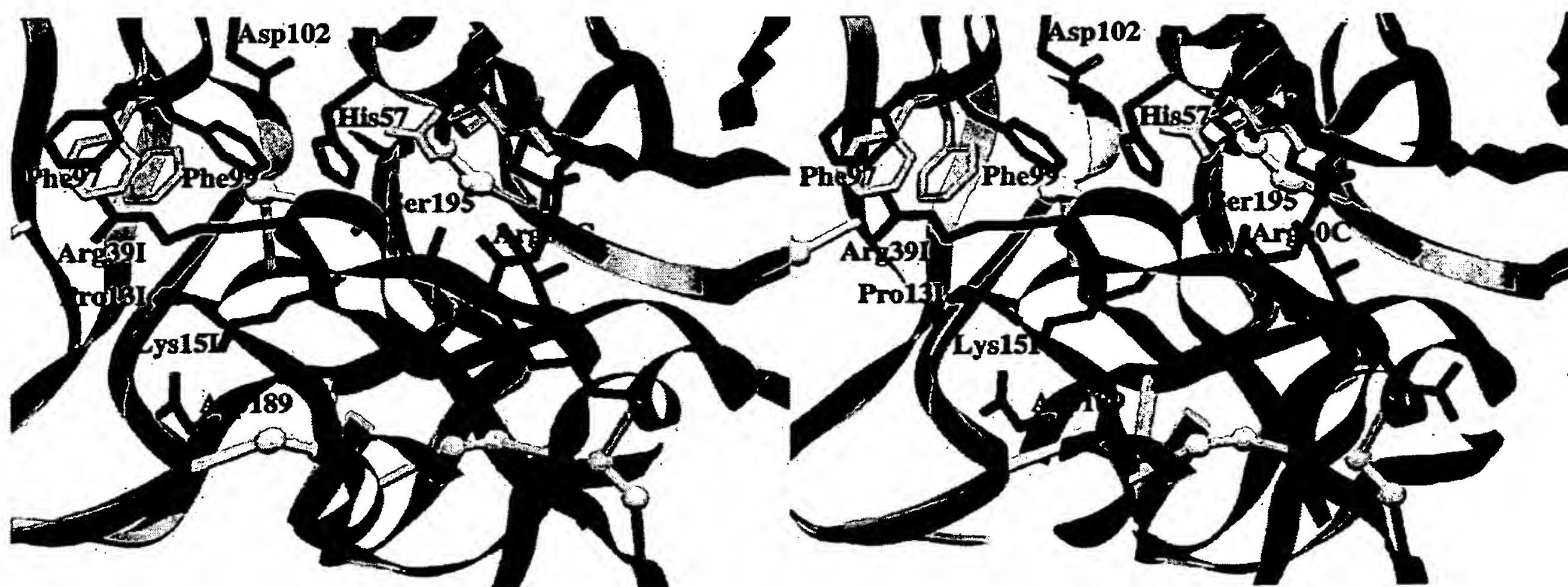


FIG. 5. Section of the complex formed between BPTI (blue color) and MT-SP1 (orange) around the contact interface, superimposed with parts of the Bz-MT-SP1 structure (red). The active site residues and Asp<sup>189</sup> (green), the benzyl groups of Phe<sup>97</sup> and Phe<sup>99</sup>, and Asp<sup>60</sup>(B) and Arg<sup>60</sup>(C) from the 60 loop of both MT-SP1 structures, as well as Pro<sup>13</sup>(I), Lys<sup>15</sup>(I), Ala<sup>16</sup>(I), Arg<sup>17</sup>(I), Ile<sup>18</sup>(I), and Arg<sup>39</sup>(I) of BPTI are shown as stick models. The large rotation of the Phe<sup>99</sup> side chain and the adaptation of the 60 loop upon binding of BPTI are clearly visible. The view is slightly tilted in comparison with Fig. 1. Figure was made with SETOR (38).

Bz-MT-SP1, in striking contrast to the equivalent Arg<sup>97</sup> residue in enteropeptidase (19). The hydrophobic S1'/S3' cavity to the east of Ser<sup>195</sup> is centered on the Cys<sup>42</sup>-Cys<sup>58</sup> disulfide bridge, and bordered by the isobutyl side chain of Ile<sup>41</sup> and the aromatic side chains of Tyr<sup>60</sup>(G) and Trp<sup>64</sup>. These positions can thus be preferentially filled with large hydrophobic P1' or P3' side chains, but also the adjacent S2' pocket, lined by Gln<sup>192</sup> and His<sup>143</sup>, gives space for hydrophobic residues (Fig. 4).

**Interaction with Kunitz-type Inhibitors**—Because the prototypic Kunitz-type inhibitor BPTI is closely related to the first Kunitz domain of HAI-1 at the sequence level, we crystallized the complex of BPTI with MT-SP1 and solved its structure at 2.9-Å resolution (Figs. 5 and 6). BPTI docks into the concave substrate binding surface of MT-SP1 via the reactive site loop (Thr<sup>11</sup>(I) to Ile<sup>18</sup>(I), with the Lys<sup>15</sup>(I) ↓ Ala<sup>16</sup>(I) scissile bond) and the secondary binding segment (Gly<sup>36</sup>(I) to Arg<sup>39</sup>(I)), similar to the prototypical trypsin complex (23). The reactive site loop of BPTI runs anti-parallel to MT-SP1 segments Ser<sup>214</sup>–Asp<sup>217</sup> and His<sup>40</sup>–Ile<sup>41</sup>, with the P1-Lys<sup>15</sup>(I) side chain extending into the S1 pocket and interacting via its ε-ammonium group with the Asp<sup>189</sup> carboxylate. Because of the main chain kink at P3-Pro<sup>13</sup>(I), which is typical for Kunitz-type inhibitors, the P3 pyrrolidine ring nestles into the S4 depression.

The Phe<sup>99</sup> side chain, if positioned as observed in Bz-MT-SP1, would clash with the Cys<sup>14</sup>(I)-Cys<sup>38</sup>(I) disulfide bridge. In the BPTI complex, this clash is avoided by a rotation of the Phe<sup>99</sup> benzyl group away from this site, in this way enlarging the S2 cavity and making it accessible for the disulfide bridge (Fig. 5). The Arg<sup>39</sup>(I) side chain extends toward the 99 loop of MT-SP1, with its guanidyl group stacking between the Phe<sup>97</sup> and the Phe<sup>99</sup> benzyl moieties and hydrogen bonding to the Phe<sup>97</sup> carbonyl oxygen much more favorably than in the BPTI-trypsin complex, because of the additional shielding by the Phe<sup>97</sup> side chain. On the primed side of the reactive center loop, the P1'-Ala<sup>16</sup>(I) and the P3'-Ile<sup>18</sup>(I) side chains form a hydrophobic knob that interacts with the hydrophobic S1'/S3' cavity.

The characteristic protruding 60 loop of MT-SP1 provides a large extra surface to make a number of favorable new contacts with BPTI, and hence with any other canonically bound Kunitz domain (Fig. 5). A comparison with Bz-MT-SP1 shows that BPTI induces some conformational changes in the 60 loop, to allow for instance the formation of a salt bridge between

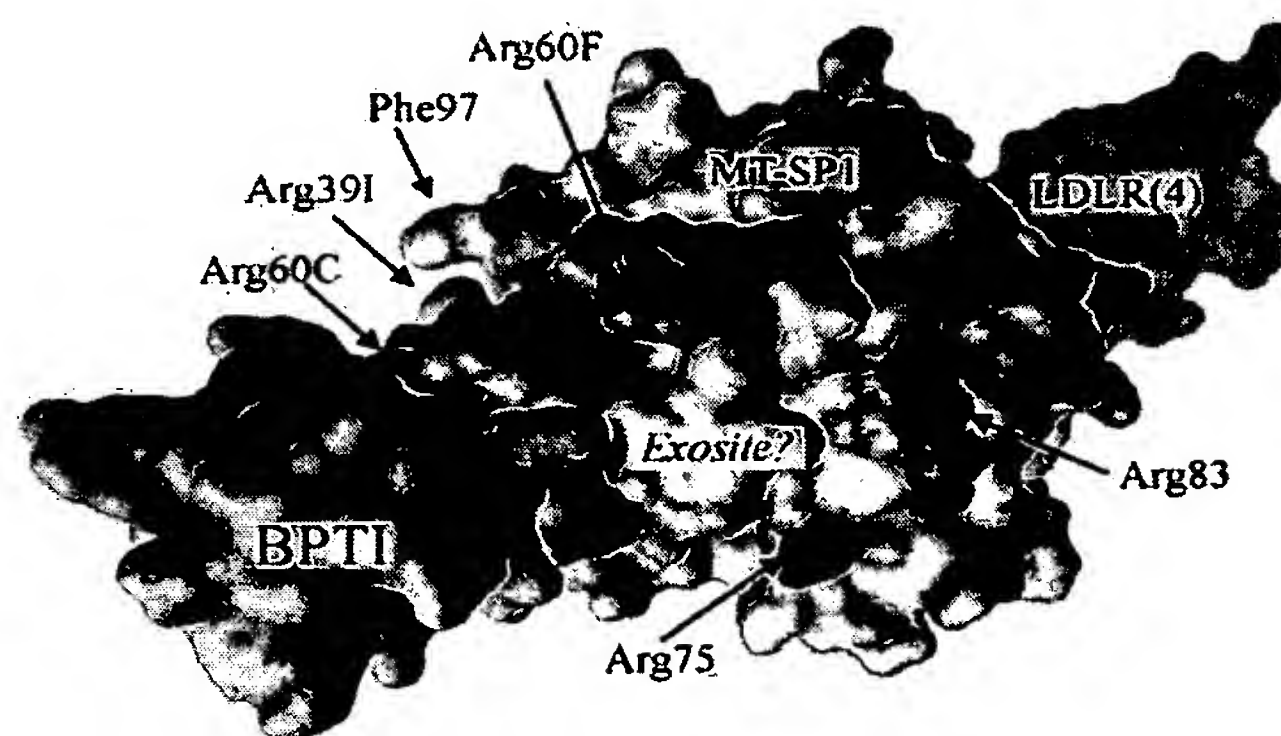


FIG. 6. Docking of BPTI (yellow) and of the (covalently linked) fourth LDLR domain (green) to opposite sites of the MT-SP1 catalytic domain. Solid surface representations, with the colors of the MT-SP1 component showing negative (red) and positive (blue) surface potential (as above). Some surface residues and the thrombin-like anion binding exosite I are indicated by labels. The complex is rotated about a vertical axis for 90° compared with the standard orientation in Fig. 1. Figure was made with WEBLABVIEWER (www.msi.com).

Asp<sup>60</sup>(B) (MT-SP1) and Arg<sup>20</sup>(I), and of charged hydrogen bonds between Arg<sup>60</sup>(C) and the carbonyl groups of Lys<sup>41</sup>(I) and Asn<sup>44</sup>(I). The well defined 60 loop of MT-SP1 thus exhibits some capability to adapt to the rigid Kunitz-type inhibitor. In this respect it strongly differs from the also exposed but much more rigid thrombin 60 loop, which delimits the S2 cavity and normally prevents access of bulkier substrates and protein inhibitors (30). The 60 loop of MT-SP1, in contrast, does not impair BPTI binding, but instead strengthens it by forming a number of additional favorable interactions (Figs. 5 and 6).

#### DISCUSSION

The catalytic domain of human MT-SP1 reveals the overall fold of a (chymo)trypsin-like serine proteinase, but displays unique properties such as the hydrophobic/acidic S2/S4 subsites and an exposed 60 loop, which considerably affect its substrate recognition and binding properties. The MT-SP1 polypeptide fold deviates notably from that of the serine proteinase trypsin, however, which exhibits six novel surface



loops through which the tryptase monomers form the intermolecular interactions that stabilize the unusual tetramer structure (31). Thus, MT-SP1 is not closely related to tryptase, and the name "matriptase" is therefore potentially confusing. However, because enterokinase and hepsin are members of the "MT-SP" family that were discovered and described before matriptase, assignment of the name MT-SP1 to matriptase may also produce controversy. One potential solution to these nomenclature issues would be to adopt either the "MT-SP" or the TTSP (3) terminology that was suggested previously by others, and to assign numbers to individual family members based on their date of discovery.

MT-SP1 can cleave selected synthetic substrates as effectively as trypsin, but exhibits a significantly more restricted specificity than trypsin (10). This may reflect, on the one hand, the near identity of the S1 specificity pockets of these two enzymes, presumably allowing both tight binding of substrates and optimal presentation of their scissile peptide bonds to the enzyme active site, and, on the other hand, clear structural distinctions between the extended binding subsites of the two proteinases. With respect to shape and chemical composition, the S1 pocket of MT-SP1 (and trypsin) provides good complementarity for Lys as well as arginine P1 residues. The efficient accommodation of P1-Lys residues, which has been demonstrated experimentally (5, 10), is facilitated by Ser<sup>190</sup>, whose side chain provides an additional hydrogen bond acceptor to stabilize the buried  $\alpha$ -ammonium group. P1-Arg residues are accommodated equally well, because of the overall better space filling by the guanidinium group (30).

The hydrophobic S2 groove of MT-SP1 is shaped to accommodate small to medium-sized hydrophobic side chains of P2-L-amino acids, with a wide hydrophobic exit toward the bulk solvent that would expose longer and more polar side chains. In addition, the rotation of the Phe<sup>99</sup> benzyl group upon BPTI binding suggests that the S2 subsite of MT-SP1 is not rigid. This observation is consistent with experimental findings from positional scanning (10), which indicated that MT-SP1 accepts a broad range of amino acids in the P2 position.

Craik and co-workers (10) have reported previously that MT-SP1 exhibits a strong preference for peptide substrates that contain Arg or Lys at P4 position, and a certain preference for either Gln or basic residues at P3. This interesting specificity may be mediated by electrostatic interactions with the acidic side chains of Asp-217 and/or Asp-96 (see Fig. 4), which could favorably pre-orient specific basic peptide substrates as they approach the enzyme active site cleft. Because canonically binding substrates align antiparallel to the Ser<sup>214</sup>-Asp<sup>217</sup> segment in an extended conformation, the side chains of P4-Arg or Lys residues presumably extend toward the 99 loop with their guanidyl or ammonium groups not reaching the Asp<sup>96</sup> carboxylate directly but forming a charged hydrogen bond with Phe<sup>97</sup> O. In addition, they can favorably interact with the overall negative potential created by the 96-, 97-, and 98-carbonyls, in full agreement with the preference of MT-SP1 for basic P4 residues. This MT-SP1 S4 subsite is reminiscent of the corresponding region of coagulation factor Xa, which has also been shown to function as a cation binding site (32). The Arg<sup>39</sup>(I) side chain of BPTI (see Fig. 5), although provided by the secondary binding segment of the inhibitor, is a nice example of such interactions; in the BPTI-MT-SP1 complex, the hydrogen bond between the Arg<sup>39</sup>(I) guanidyl and the Phe<sup>97</sup> O is shielded from solvent by the unique Phe<sup>97</sup> benzyl side chain (see Figs. 4 and 5), which would significantly strengthen these interactions.

In canonically bound peptide substrates, the side chain of a P3 residue projects out of the active site cleft, where it can

hydrogen-bond the carboxamide group of Gln<sup>192</sup>. It is also possible, however, that a bound substrate could adopt a "kinked" conformation at the P3 position, as seen for Kunitz-type inhibitors (Fig. 5). In this alternative conformation, the P3 side chain extends into the S4 subsite, *i.e.* toward the 99 loop. To allow the guanidyl or the ammonium group of a P3-Arg or Lys residue, bound to MT-SP1 in the alternative conformation, to form direct hydrogen bonds with Phe<sup>97</sup> O, the substrate main chain around P3 would have to rotate and thereby weaken considerably the inter-main chain hydrogen bonds to Gly<sup>216</sup>. In either conformation, however, a basic P3 side chain would interact favorably with the negative potential of the MT-SP1 S4 pocket (Fig. 4). Long range electrostatic interactions between side chain charges and complementary surface potentials are often found in protein-protein complexes. For instance, in the thrombin complexes with hirudin and other protein inhibitors, the removal of the charges of acidic residues not involved in direct salt bridges has been shown to affect the electrostatic interactions strongly (33, 34). The relatively low probability of the simultaneous occurrence of Arg/Lys residues at P3 and P4 in good MT-SP1 substrates (10) would then be explained by mutual charge compensation and exclusion from the same (S4) site.

The specificity of MT-SP1 differs substantially from that of trypsin, because MT-SP1 does not indiscriminately cleave peptide substrates at accessible Lys or Arg residues, but instead requires recognition of additional residues surrounding the scissile peptide bond. This requirement for recognition of an extended primary sequence for efficient catalysis of substrates suggests that MT-SP1 is a relatively specific proteinase that may play a regulatory role (5, 6). Recognition of an extended primary sequence appears to be also required for efficient cleavage of macromolecular substrates by MT-SP1. Efficient auto-activation of MT-SP1 entails recognition and cleavage of an Arg-Gln-Ala-Arg P4-P1 target sequence. MT-SP1 can also efficiently activate the proteinase-activated receptor-2 (PAR2), sc-uPA (10), and the hepatocyte growth factor/scatter factor (11). These extracellular surface-localized proteins display the P4 to P1 target sequences Ser-Lys-Gly-Arg, Pro-Arg-Phe-Lys, and Lys-Gln-Gly-Arg, respectively, which match closely the MT-SP1 cleavage specificity requirements observed for small peptidic substrates (10). Another indication of the substrate specificity of MT-SP1 is that the enzyme does not activate proteins closely related to these substrates, such as PAR-1, PAR-3, PAR-4, and plasminogen, that do not display target sequences matching the extended MT-SP1 specificity near the scissile bond.

Because MT-SP1 has been found co-localized with sc-uPA in several cell types, it has been suggested that MT-SP1 may be a physiologically relevant activator of this proteinase zymogen (10). uPA plays an important role in angiogenesis and/or tumor progression and can also activate both plasminogen and matrix metalloproteinases (35). The latter, in turn, play important roles in matrix degradation and re-modeling, events that are required both during angiogenesis and tumor invasion and metastasis. In addition, MT-SP1 can directly activate hepatocyte growth factor, a protein that promotes cell growth as well as angiogenesis; therefore, it may play both direct and indirect roles in cell growth and migration and, when improperly regulated, may contribute to tumor angiogenesis, growth, and progression.

In normal tissues, the proteolytic activity of proteinases is carefully controlled localizing their action in time and space. As mentioned above, HAI-1, a type-II transmembrane protein containing two Kunitz-type domains, appears to be the primary physiological inhibitor of MT-SP1 (12). Based on the structure

of our BPTI-MT-SP1 complex, we have modeled the complex between MT-SP1 and the first Kunitz domain of HAI-1, which is 36% identical to BPTI and appears more likely to exhibit high affinity for the MT-SP1 active site than the second HAI-1 Kunitz domain. This model suggests that a large number of favorable interactions could form between the first HAI-1 Kunitz domain and MT-SP1, both between the Gly<sup>12</sup>(I)-Arg<sup>13</sup>(I)-Cys<sup>14</sup>(I)-Arg<sup>15</sup>(I) ↓ Gly<sup>16</sup>(I)-Ser<sup>17</sup>(I)-Phe<sup>18</sup>(I) reactive site loop (using the BPTI nomenclature, with Arg<sup>15</sup>(I) ↓ Gly<sup>16</sup>(I) representing the scissile bond) and the active site cleft, and also at secondary interaction sites such as made by the 60 loop. In the primed side, the first HAI-1 Kunitz domain possesses an uncommon P3'-Phe<sup>18</sup>(I), which could, because of the small P1'-Gly residue, be nicely packed in the large hydrophobic S1'/S3' pocket of MT-SP1 (Fig. 4). The side chain of the P3-equivalent Arg<sup>13</sup>(I) is expected to extend into the S4 subsite, where its guanidyl group would make favorable electrostatic interactions with the 99 loop carbonyls, similar to those observed for Arg<sup>39</sup>(I) in the BPTI complex (see Fig. 5).

The putative Gly<sup>12</sup>(I)-Leu<sup>13</sup>(I)-Cys<sup>14</sup>(I)-Lys<sup>15</sup>(I) ↓ Glu<sup>16</sup>(I)-Ser<sup>17</sup>(I)-Ile<sup>18</sup>(I) reactive site loop of the second Kunitz domain of HAI-1, in contrast, does not match the reported substrate specificity of MT-SP1. The second HAI-1 domain does, however, possess a number of negatively charged residues in its C-terminal α-helix that could form favorable electrostatic interactions with basic surface residues of MT-SP1 that map in or near the region corresponding to the anion binding exosite I of thrombin (such as Arg<sup>75</sup>, Arg<sup>83</sup>, Arg<sup>85</sup>, Lys<sup>110</sup>; see Fig. 6). Such an additional exosite binding of the second HAI-1 domain accompanying the interaction of the first domain with the MT-SP1 active site would considerably increase the affinity and specificity of HAI-1 for MT-SP1. A similar cooperative action with consequent increase in affinity and specificity has been previously observed in the complex between thrombin and orithodorin, a two-domain Kunitz-type inhibitor derived from the blood-sucking tick *Ornithodoros moubata* (36), where the first and the second Kunitz domains interact (noncanonically, however) with the active site and electrostatically with exosite I, respectively, of thrombin.

The role of the four LDLR domains that precede the catalytic domain remains unclear, but they have been implicated in mediating interactions with other membrane or membrane-associated proteins (5, 6). The LDLR (4) domain of MT-SP1 was modeled (Fig. 6) based on the structure of the fifth low density lipoprotein class A binding domain of LDLR (37), which shares a conserved disulfide bonding pattern and a 45% amino acid identity with LDLR (4). In full-length MT-SP1, the last cysteine residue (Cys<sup>604</sup>(g)) of LDLR (4) is separated by only one amino acid (Asp<sup>605</sup>(g)) from the first residue (Cys<sup>606</sup>(g) = Cys<sup>1</sup>) of the catalytic domain, which, in turn, forms an intradomain disulfide bridge with Cys<sup>122</sup>. This implicates close proximity of both modules. A careful inspection of the electrostatic potentials suggests a distinct rotational orientation of LDLR (4) and the catalytic domain relative to each other, which would create both favorable interactions between electrostatic potentials of the two domains and a good steric fit of the two complementary surfaces. Our docking experiment predicts four interdomain salt bridges (Glu<sup>593</sup>(g)-Arg<sup>206</sup>, Lys<sup>592</sup>(g)-Asp<sup>125</sup>, Asp<sup>605</sup>(g)-Arg<sup>119</sup>, and Asp<sup>605</sup>(g)-Lys<sup>602</sup>(g)) and charged hydrogen bonds from Lys<sup>239</sup> and Arg<sup>235</sup> to the 589(g)-590(g) backbone. The LDLR (4) domain of MT-SP1 carries all four acidic side chains (of Asp<sup>590</sup>(g), Asp<sup>594</sup>(g), Asp<sup>600</sup>(g), and Glu<sup>601</sup>(g)) engaged in calcium coordination in the fifth low density lipoprotein-A domain of LDLR (37). The suggested association of LDLR (4) with the catalytic domain does not directly involve this putative calcium binding site, which would be located close to the inter-

face, however. The location of the LDLR (4) domain on the proteinase surface opposite to the active site cleft allows unrestricted substrate or inhibitor binding to MT-SP1.

It has only recently been appreciated that TTSPs represent an important, emerging subfamily of the (chymo)trypsin enzyme family. With the exception of enterokinase, an important digestive enzyme with a unique substrate specificity, the biological roles of individual members of this recently discovered enzyme subfamily have not yet been unambiguously established. The first review of this new class was published earlier this year and discussed seven human TTSPs (3). As described by these authors, the majority of these human proteinases have been associated either with tumor cells and/or cell growth. Because both cell growth and tumor progression (including angiogenesis) are expected to require localized activation of growth factors and degradation and/or remodeling of the extracellular matrix, it seems likely that these processes require cell associated, proteolytic activities. Other membrane-associated proteinases such as the disintegration and metalloproteinase (ADAM), matrix metalloproteinases, and uPA have previously been implicated in these processes, and it is conceivable that specific members of the TTSPs will also contribute to these key processes. The development of highly potent and selective inhibitors of individual TTSPs, therefore, may represent an exciting new strategy to discover compounds with anti-angiogenic or antitumor activities.

MT-SP1 has been implicated in the progression of prostate cancer in a rat model, and our solution of the high resolution structures of the proteinase domain, in complex with benzamidine or BPTI, is an important first step toward the design of potent selective inhibitors. Because the proteinase active site is fully solvent-exposed in the Bz-MT-SP1 crystals, as demonstrated by initial soaking experiments (data not shown), small molecule inhibitors can be soaked into these crystals, thereby facilitating rapid progress in structure-based drug design efforts. In addition, the structures reported here can be used to model the structure of other cancer-associated type II transmembrane serine proteinases, and may therefore facilitate efforts to find potent, specific inhibitors of those important proteinases as well. First homology-based inhibitor design studies have been undertaken by Enyedy *et al.* (16), based on a thrombin model, which led to the development of selective bis-benzamidine inhibitors. This work showed that there is sufficient structural difference between MT-SP1 and closely related serine proteinases for the development of potent and selective inhibitors. Small molecule inhibitors of TTSPs can be used not only to discover the biological and pathological roles of specific members of this intriguing new enzyme subfamily but, because of the strong association of these enzymes with tumor cells, can also be tested as potential lead anticancer compounds.

**Acknowledgments**—We thank George Vlasuk for many useful discussions, Trish Nguyen for expert technical assistance, and John L. Richardson for English language advice.

#### REFERENCES

- Mignatti, P., and Rifkin, D. B. (1993) *Physiol. Rev.* **73**, 161–195
- Kute, T. E., Grondahlansen, J., Shao, S. M., Long, R., Russell, G., and Brunner, N. (1998) *Breast Cancer Res. Treat.* **47**, 9–16
- Hooper, J. D., Clements, J. A., Quigley, J. P., and Antalis, T. M. (2001) *J. Biol. Chem.* **276**, 857–860
- Shi, Y. E., Torri, J., Yieh, L., Wellstein, A., Lippman, M. E., and Dickson, R. B. (1993) *Cancer Res.* **53**, 1409–1415
- Lin, C. Y., Anders, J., Johnson, M., Sang, Q. A., and Dickson, R. B. (1999) *J. Biol. Chem.* **274**, 18231–18236
- Takeuchi, T., Shuman, M. A., and Craik, C. S. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 11054–11061
- Benaud, C., Dickson, R. B., and Lin, C. Y. (2001) *Eur. J. Biochem.* **268**, 1439–1447
- Oberst, M., Anders, J., Xie, B., Singh, B., Ossandon, M., Johnson, M., Dickson, R. B., and Lin, C. Y. (2001) *Am. J. Pathol.* **158**, 1301–1311
- Kim, M. G., Chen, C., Lyu, M. S., Cho, E. G., Park, D., Kozak, C., and Schwartz,



- R. H. (1999) *Immunogenetics* **49**, 420–428
10. Takeuchi, T., Harris, J. L., Huang, W., Yan, K. W., Coughlin, S. R., and Craik, C. S. (2000) *J. Biol. Chem.* **275**, 26333–26342
  11. Lee, S. L., Dickson, R. B., and Lin, C. Y. (2000) *J. Biol. Chem.* **275**, 36720–36725
  12. Lin, C. Y., Anders, J., Johnson, M., and Dickson, R. B. (1999) *J. Biol. Chem.* **274**, 18237–18242
  13. Shimomura, T., Denda, K., Kitamura, A., Kawaguchi, T., Kito, M., Kondo, J., Kagaya, S., Qin, L., Takata, H., Miyazawa, K., and Kitamura, N. (1997) *J. Biol. Chem.* **272**, 6370–6376
  14. Shimomura, T., Denda, K., Kawaguchi, T., Matsumoto, K., Miyazawa, K., and Kitamura, N. (1999) *J. Biochem. (Tokyo)* **126**, 821–828
  15. Qin, L., Denda, K., Shimomura, T., Kawaguchi, T., and Kitamura, N. (1998) *FEBS Lett.* **436**, 111–114
  16. Enyedy, I. J., Lee, S. L., Kuo, A. H., Dickson, R. B., Lin, C. Y., and Wang, S. (2001) *J. Med. Chem.* **44**, 1349–1355
  17. Collaborative Computational Project No. 4. (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760–763
  18. Navaza, J. (1994) *Acta Crystallogr. Sect. A* **50**, 157–163
  19. Lu, D., Futterer, K., Korolev, S., Zheng, X., Tan, K., Waksman, G., and Sadler, J. E. (1999) *J. Mol. Biol.* **292**, 361–373
  20. Turk, D. (1992) *Further Development of a Program for Molecular Graphics and for the Manipulation of Electron Densities and Its Application to Various Protein Structure Determinations*. Ph.D. thesis, Technische Universität, München
  21. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54**, 905–921
  22. Engh, R., and Huber, R. (1991) *Acta Crystallogr. Sect. A* **47**, 392–400
  23. Huber, R., Kukla, D., Bode, W., Schwager, P., Bartels, K., Deisenhofer, J., and Steigemann, W. (1974) *J. Mol. Biol.* **89**, 73–101
  24. Nienaber, V., Wang, J. Y., Davidson, D., and Henkin, J. (2000) *J. Biol. Chem.* **275**, 7239–7248
  25. Bode, W., and Schwager, P. (1975) *J. Mol. Biol.* **98**, 693–717
  26. Birktoft, J. J., and Blow, D. M. (1972) *J. Mol. Biol.* **68**, 187–240
  27. Bode, W., Turk, D., and Karshikov, A. (1992) *Protein Sci.* **1**, 426–471
  28. Bode, W., Schwager, P., and Huber, R. (1978) *J. Mol. Biol.* **118**, 99–112
  29. Madison, E. L., Kobe, A., Gething, M. J., Sambrook, J. F., and Goldsmith, E. J. (1993) *Science* **262**, 419–421
  30. Bode, W., and Huber, R. (1992) *Eur. J. Biochem.* **204**, 433–451
  31. Pereira, P. J., Bergner, A., Macedo-Ribeiro, S., Huber, R., Matschiner, G., Fritz, H., Sommerhoff, C. P., and Bode, W. (1998) *Nature* **392**, 306–311
  32. Stubbs, M. T., Huber, R., and Bode, W. (1995) *FEBS Lett.* **375**, 103–107
  33. Karshikov, A., Bode, W., Tulinsky, A., and Stone, S. R. (1992) *Protein Sci.* **1**, 727–735
  34. Stubbs, M. T., and Bode, W. (1995) *Trends Biochem. Sci.* **20**, 131
  35. Andreasen, P. A., Egelund, R., and Petersen, H. H. (2000) *Cell. Mol. Life Sci.* **57**, 25–40
  36. van de Locht, A., Bode, W., Huber, R., Le Bonniec, B. F., Stone, S. R., Esmon, C. T., and Stubbs, M. T. (1997) *EMBO J.* **16**, 2977–2984
  37. Fass, D., Blacklow, S., Kim, P. S., and Berger, J. M. (1997) *Nature* **388**, 691–693
  38. Evans, S. V. (1990) *J. Mol. Graph.* **11**, 134–138
  39. Blevins, R. A., and Tulinsky, A. (1985) *J. Biol. Chem.* **260**, 4264–4275
  40. Kitamoto, Y., Veile, R. A., Donis-Keller, H., and Sadler, J. E. (1995) *Biochemistry* **34**, 4562–4568
  41. Barton, G. J. (1993) *Protein Eng.* **6**, 37–40
  42. Nicholls, A., Bharadwaj, R., and Honig, B. (1993) *Biophys. J.* **64**, A166
  43. Leslie, A. (1991) in *Crystal Computing V* (Moras, D., Podjarny, A. D., and Thierry, J. C., eds) pp. 27–38, Oxford University Press, Oxford

# The refined 1.9-Å X-ray crystal structure of D-Phe-Pro-Arg chloromethylketone-inhibited human $\alpha$ -thrombin: Structure analysis, overall structure, electrostatic properties, detailed active-site geometry, and structure–function relationships



WOLFRAM BODE, DUSAN TURK,<sup>1</sup> AND ANDREJ KARSHIKOV

Max-Planck-Institut für Biochemie, D-8033 Martinsried, Germany

(RECEIVED October 29, 1991; REVISED MANUSCRIPT RECEIVED November 18, 1991)

## Abstract

Thrombin is a multifunctional serine proteinase that plays a key role in coagulation while exhibiting several other key cellular bioregulatory functions. The X-ray crystal structure of human  $\alpha$ -thrombin was determined in its complex with the specific thrombin inhibitor D-Phe-Pro-Arg chloromethylketone (PPACK) using Patterson search methods and a search model derived from trypsinlike proteinases of known spatial structure (Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S.R., & Hofsteenge, J., 1989, *EMBO J.* 8, 3467–3475). The crystallographic refinement of the PPACK–thrombin model has now been completed at an *R* value of 0.156 (8 to 1.92 Å); in particular, the amino- and the carboxy-termini of the thrombin A-chain are now defined and all side-chain atoms localized; only proline 37 was found to be in a cis-peptidyl conformation.

The thrombin B-chain exhibits the characteristic polypeptide fold of trypsinlike serine proteinases; 195 residues occupy topologically equivalent positions with residues in bovine trypsin and 190 with those in bovine chymotrypsin with a root-mean-square (r.m.s.) deviation of 0.8 Å for their  $\alpha$ -carbon atoms. Most of the inserted residues constitute novel surface loops. A chymotrypsinogen numbering is suggested for thrombin based on the topological equivalences. The thrombin A-chain is arranged in a boomeranglike shape against the B-chain globule opposite to the active site; it resembles somewhat the propeptide of chymotrypsin(ogen) and is similarly not involved in substrate and inhibitor binding.

Thrombin possesses an exceptionally large proportion of charged residues. The negatively and positively charged residues are not distributed uniformly over the whole molecule, but are clustered to form a sandwichlike electrostatic potential; in particular, two extended patches of mainly positively charged residues occur close to the carboxy-terminal B-chain helix (forming the presumed heparin-binding site) and on the surface of loop segment 70–80 (the fibrin[ogen] secondary binding exosite), respectively; the negatively charged residues are more clustered in the ringlike region between both poles, particularly around the active site. Several of the charged residues are involved in salt bridges; most are on the surface, but 10 charged protein groups form completely buried salt bridges and clusters. These electrostatic interactions play a particularly important role in the intrachain stabilization of the A-chain, in the coherence between the A- and the B-chain, and in the surface structure of the fibrin(ogen) secondary binding exosite (loop segment 67–80).

The most remarkable feature at the thrombin surface is the prominent canyonlike active-site cleft mainly shaped by two characteristic insertion loops around Trp 60D and Trp 148. The deep and narrow active-site cleft in general explains the narrow specificity of thrombin for distinct macromolecular substrates and inhibitors. Comparisons with other crystal structures of human and bovine thrombin recently determined using this PPACK–thrombin model indicate that the first loop around Trp 60D is relatively rigid, whereas the opposite loop around Trp 148 can attain different conformations depending on complexation state and crystalline environment.

The active-site residues and the entrance to the specificity pocket are partially occluded in thrombin (much more than in the other serine proteinases) by this distinctive Trp 60D loop. The specificity pocket of thrombin resem-

Reprint requests to: Wolfram Bode, Max-Planck-Institut für Biochemie, D-8033 Martinsried, Germany.

<sup>1</sup> On leave from the Chemical Institut Boris Kidric, Ljubljana, Slovenia.



bles that of bovine trypsin but is designed to prefer arginine over lysine residues at P1. D-Phe 11 and Pro 21 of the bound PPACK inhibitor fit neatly to a novel hydrophobic cleft (the aryl-binding site) and to the cavitylike hydrophobic S2 subsite; the D-configuration of Phe 11 is beneficial for binding as it allows the PPACK amino-terminus to form hydrogen bonds to Gly 216 in addition. Some small arginine and benzamidine-derived synthetic inhibitors owe their particularly high thrombin specificity and affinity to their exceptional steric fit to these novel hydrophobic cavities close to the thrombin active site (Bode, W., Turk, D., & Stürzebecher, J., 1990, *Eur. J. Biochem.* 193, 175–182).

The active-site cleft levels off in the primed direction and continues over the molecular surface of the thrombin loop Lys 70–Glu 80 (itself structurally similar to the calcium loop in trypsin, with the distal nitrogen of Lys 70 replacing the calcium of trypsin). This site of a strong positive electrostatic surface potential probably represents the secondary site for interaction with the  $\alpha$ -chain of fibrinogen and fibrin (the fibrin[ogen] secondary binding exosite) and accommodates the carboxy-terminal acidic tail part of hirudin (Rydel, T.J., Ravichandran, K.G., Tulinsky, A., Bode, W., Huber, R., Roitsch, C., & Fenton, J.W., II, 1990, *Science* 249, 277–280). Segment Arg 187–Gly 188–Asp 189, which could represent a thrombin adhesion site for cellular interactions with platelets, fibroblasts, and endothelial cells, is mainly buried in  $\alpha$ -thrombin; its adhesive role would thus appear to require some prior unfolding.

Most of the well-characterized sites of proteolytic cleavage leading to the degradation products  $\beta$ -,  $\gamma$ -, and  $\epsilon$ -thrombin of diminished or lost clotting activity are situated in exposed mobile loops of  $\alpha$ -thrombin. None of these segments is in a canonical conformation that would allow association with the substrate-binding site of a cleaving serine proteinase without large conformational changes. The cleavage of the Arg 77A–Asn 78 scissile peptide bond (leading to  $\beta$ -thrombin) presumably results in the unfolding of the 70–80 loop, exposing the salt bridge-connected residues buried in  $\alpha$ -thrombin to the solvent; the concomitant disruption of the surface of the fibrin(ogen) secondary binding exosite would explain the loss of binding capacity (and thus catalytic activity) toward fibrinogen. The Arg 67–Ile 68 peptide bond is completely buried beneath this 70–80 loop surface and thus only susceptible to an attacking proteinase after prior cleavage and exposure of this loop, as observed.

**Keywords:** electrostatic interactions; protein crystallography; serine proteinase; thrombin; thrombosis

$\alpha$ -Thrombin<sup>1</sup> is a glycosylated trypsinlike serine proteinase generated in the penultimate step of the blood coagulation cascade from the circulating plasma protein prothrombin. Upon autocatalytic and Factor Xa cleavage the functional two-chain molecule  $\alpha$ -thrombin is generated. In the case of the human species this  $\alpha$ -thrombin consists of the 36-residue A-chain and the 259-residue B-chain (Butkowski et al., 1977; Thompson et al., 1977; Degen et al., 1983). The two chains are connected covalently by a disulfide bridge. The B-chain has been shown to be homologous to the catalytic domains of other pancreatic and coagulation/fibrinolytic trypsinlike proteinases (Jackson & Nemerson, 1980). Upon further autolytic or proteolytic cleavage, more species (in particular  $\beta$ - and  $\gamma$ -thrombin) are generated that retain some activity against small synthetic substrates but have lost most or all clotting activity (Lundblad et al., 1979; Elion et al., 1986; Hofsteenge et al., 1988).

Thrombin is a multifunctional protein; it plays a central role in thrombosis and hemostasis but is also impli-

cated in wound healing and various disease processes (see Fenton, 1988).  $\alpha$ -Thrombin converts fibrinogen into fibrin, which consequently aggregates and forms the interconnecting network of thrombi. Furthermore, thrombin is able to activate several coagulation and plasma factors, such as Factors V, VIII, XIII, and protein C. The complex of thrombin with thrombomodulin exhibits enhanced reactivity toward protein C (Esmon et al., 1982) but has lost all other procoagulant activity (Esmon et al., 1983; Hofsteenge & Stone, 1987). Thrombin is effectively inhibited by only a very few endogenous protein inhibitors, such as  $\alpha_2$ -macroglobulin and the serpins antithrombin III (Rosenberg & Damus, 1973), heparin cofactor II (Tollefsen et al., 1982), and protease nexin I (Cunningham & Farrell, 1986). The interaction with these serpins is considerably enhanced by the acidic glycosaminoglycan heparin (Li et al., 1976; Björk & Lindahl, 1982; Olson & Shore, 1982; Wallace et al., 1989). In addition, thrombin binds very tightly to and is selectively inhibited by hirudin, a protein isolated from the medicinal leech (Walsmann & Markwardt, 1981; Stone & Hofsteenge, 1986; Dodt et al., 1988).

Thrombin acts on a variety of cells (see Jackson & Nemerson, 1980). Some of these cellular interactions require proteolytically active forms of thrombin, whereas others also function with active-site-blocked thrombin species (see, e.g., Carney et al., 1986; Fenton, 1986; McGowan & Detwiler, 1986; Shuman, 1986; Vu et al., 1991). Thrombin induces platelet aggregation and stim-

<sup>1</sup> P1, P2, P3, etc., and P1', P2' designate substrate/inhibitor residues amino- and carboxy-terminal to the scissile peptide bond, respectively, and S1, S2, S3, etc., and S1', S2' the corresponding subsites of the cognate proteinase (Schechter & Berger, 1967). For numbering of the thrombin amino acid residues, the chymotrypsinogen nomenclature introduced by Bode et al. (1989b) is used. D-Phe-Pro-Arg chloromethylketone (PPACK) residues are designated by a suffix I, hirudin residues by a suffix H1 behind the residue number.

**Deposition:** The PPACK-thrombin have been deposited at the Brookhaven Protein Data Bank (Bernstein et al., 1977).

ulates platelet secretion. It causes mitogenesis in fibroblasts (Cunningham & Farrell, 1986; Glenn et al., 1988) and macrophagelike cells (Bar-Shavit et al., 1986), exhibits chemotactic properties (Bar-Shavit et al., 1983, 1984, 1986; Bizios et al., 1986), and binds to endothelial cells (Prescott et al., 1990; Bar-Shavit et al., 1991) and to subendothelial extracellular matrix (Bar-Shavit et al., 1989).

Thrombin cleaves a variety of proteins for which the amino acid sequence around the scissile peptide bond differs considerably (Blombäck et al., 1977; Scheraga, 1977). Its specificity is primarily trypsinlike, i.e., it cleaves behind arginine and lysine residues, with a clear preference for Arg-Xaa bonds (Liem & Scheraga, 1974; Lottenberg et al., 1983). It is, however, much more selective than trypsin toward macromolecular substrates in that it cleaves many fewer peptide bonds of permitted sequence; in fibrinogen, for example, only two Arg-Xaa bonds out of a total of 181 Arg/Lys-Xaa bonds are attacked (Blombäck et al., 1967; Hogg & Blombäck, 1978). Thrombin is particularly active toward peptidic substrates and inhibitors with a proline at the P2 position (Kettner & Shaw, 1981); a D-phenylalanine residue at P3 makes such peptides even more reactive toward thrombin. Rather reactive and selective chloromethyl, aldehyde, and boroarginine inhibitors have been developed with the peptidyl moiety D-Phe-Pro-Arg (Bajusz et al., 1978; Kettner & Shaw, 1979; Kettner et al., 1991); various synthetic thrombin inhibitors have been found and characterized based on arginine and benzamidine derivatives or on heterocyclic compounds (Okamoto et al., 1981; Stürzebecher et al., 1983; Kam et al., 1988).

In contrast to most other serine proteinases, the thrombin specificity toward fibrinogen is not determined by subsites surrounding the active residues alone. For efficient cleavage of fibrinogen (as well as of isolated A $\alpha$ -chains) the availability and integrity of a thrombin exosite quite distant from the catalytic residues is important (Hageman & Scheraga, 1974; Liem & Scheraga, 1974; Blombäck et al., 1977; Van Nispen et al., 1977; Marsh et al., 1985). This fibrin(ogen)-recognizing exosite, attributed to a rather basic surface area of thrombin around its  $\beta$ -cleavage site has also been implicated in binding of other proteins such as fibrin monomers, fibrin E-domain, and its NDSK-fragment, thrombomodulin, and hirudin, and of negatively charged cell surfaces, glass, and cation-exchange resins (Liu et al., 1979; Fenton et al., 1981, 1988; Berliner et al., 1985; Kaminski & McDonagh, 1987; Kaczmarek & McDonagh, 1988; Vali & Scheraga, 1988).

An apolar binding site close to the catalytic center has been inferred from binding and proflavine displacement studies (Thompson, 1976; Berliner & Shen, 1977). This site has been suggested as accounting for thrombin specificity with tripeptide substrates (Sonder & Fenton, 1984) and for the accommodation of large hydrophobic residues amino-terminal of the fibrinogen A $\alpha$ -cleavage site

(Meinwald et al., 1980; Marsh et al., 1985; Ni et al., 1989a,b).

Several dysfunctional genetic variants of human (pro)thrombin have been found; in three of them the site of mutation has been localized (Miyata et al., 1987; Henriksen & Mann, 1988, 1989). Recently, several recombinant thrombin mutants also have been prepared (LeBonniec & Esmon, 1991; LeBonniec et al., 1991; Wu et al., 1991).

Three-dimensional models have been proposed for the thrombin B-chain (Magnusson et al., 1975; Greer, 1981; Bing et al., 1986; Sugawara et al., 1986; Toma & Suzuki, 1989) based on the crystal structures of bovine trypsin and chymotrypsin. These models provided a general impression of the arrangement of sites involved in the various interactions of thrombin and were of some use in explaining certain structure-function relationships and in developing inhibitors (Maraganore et al., 1990). An adequate understanding of the specific and characteristic functions of thrombin and the accurate model-based design of thrombin inhibitors requires, however, knowledge of an experimentally determined structure.

We have recently communicated the X-ray crystal structure of D-Phe-Pro-Arg chloromethylketone (PPACK [Kettner & Shaw, 1979]) inhibited human  $\alpha$ -thrombin and presented a concise description of the most important features of thrombin, in particular its unique and prominent canyonlike active-site cleft (Bode et al., 1989b). This PPACK-thrombin model is now fully refined; the relatively flexible amino- and carboxy-termini of the A-chain have been determined, the locations of a few amino acid side chains have been revised, and more solvent molecules have been added in order to improve further the phases.

Using the refined coordinates of this PPACK-thrombin structure as a search and initial phasing model, the X-ray crystal structures of human  $\alpha$ -thrombin complexes of hirudin (Grütter et al., 1990; Rydel et al., 1990, 1991) and fibrinopeptide A/hirugen (Stubbs et al., 1992) and of bovine thrombin complexes (Martin et al., 1992) have been elucidated. Recently, we established the exact binding geometry of some arginine- and benzamidine-based synthetic inhibitors toward bovine thrombin (Brandstetter et al., 1992) and confirmed earlier modeling results inferred from trypsin binding (Bode et al., 1990; Chow et al., 1990; Turk et al., 1991).

In the following, the entire course of the structure analysis of PPACK-thrombin and the completely refined model are fully described; sites of particular interest, such as the A-B-chain interaction and the distinct substrate-binding subsites including the interactions with substrates and inhibitors are presented in more detail; the unique electrostatic properties of thrombin that play an important role in the molecular stability and in several molecular functions are described; the thrombin structure is compared with that of related proteinases; and finally, the various functions of thrombin are discussed with respect to its structure.

## Results and discussion

### Overall structure

The thrombin molecule can be described as a prolate ellipsoid of approximate dimensions  $45 \text{ \AA} \times 45 \text{ \AA} \times 50 \text{ \AA}$ . Its A- and B-chains are not organized in separate domains, but form a single contiguous body with a highly furrowed surface (see, e.g., Fig. 1). Similar to the other trypsinlike serine proteinases, thrombin consists essentially of two interacting six-stranded barrellike domains (linked by four transdomain "straps"), of five helical seg-

ments and one helical turn, and of various surface-located turn structures (Fig. 2; Kinemage 1). Three of these straps and three of these helices are also present in other known serine proteinase structures; the A-chain has a counterpart in chymotrypsin(ogen). The catalytic residues, in particular Ser 195, His 57, and Asp 102 (using the chymotrypsinogen nomenclature introduced by Bode et al. [1989b]; see Table 3) are located at the junction between both barrels; a prominent active-site cleft (Fig. 1) stretches perpendicular to this junction (described in more detail below).

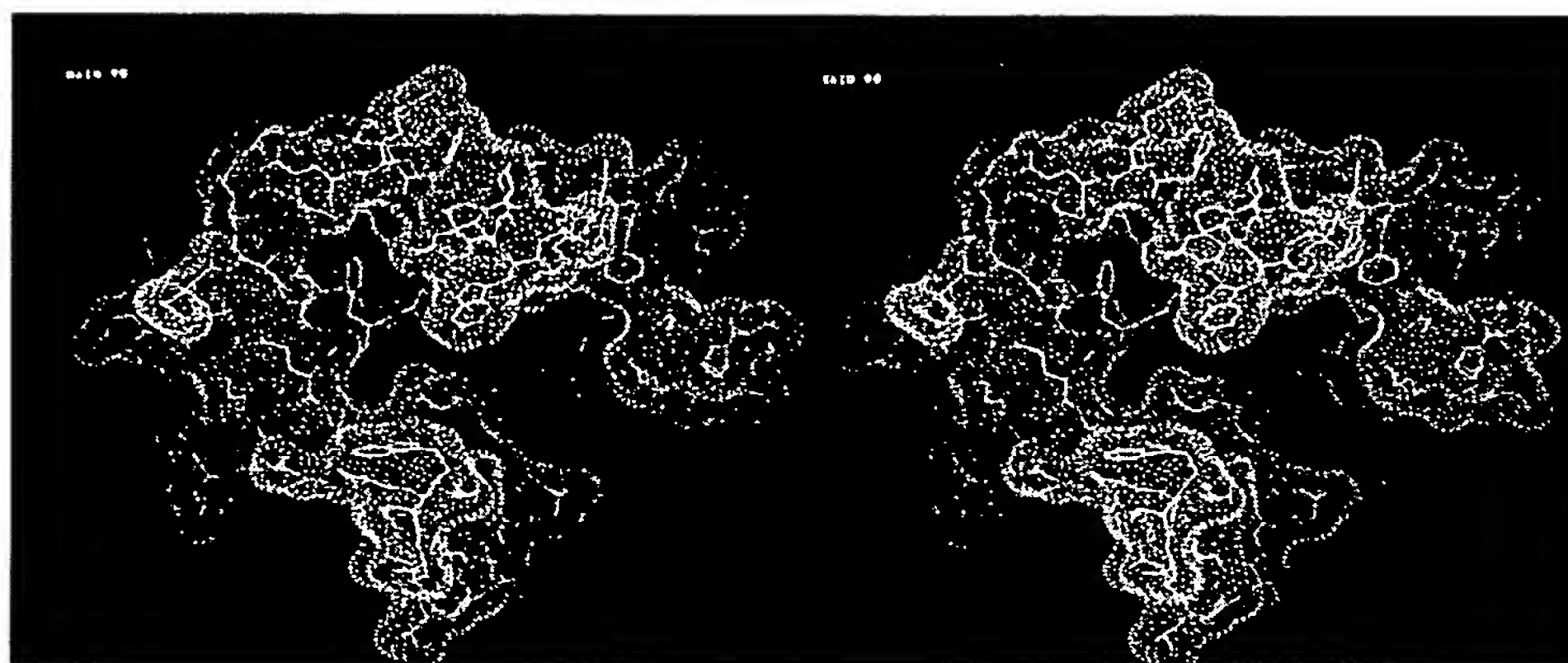


Fig. 1. Front view toward the thrombin molecule (yellow), displayed together with its Connolly dot surface (calculated with a probe of radius  $1.4 \text{ \AA}$ ). The active-site cleft runs from left to right across the molecular surface; the bound PPACK molecule (violet) is bound. Only the front parts of the thrombin molecule and of its molecular surface are displayed. A bound substrate polypeptide chain would run from left to right. The surface "hole" close to the center corresponds to the entrance to the specificity pocket. The 60 insertion loop, in particular Trp 60D, is partially occluding the active site. The fibrinogen secondary binding exosite is to the right.

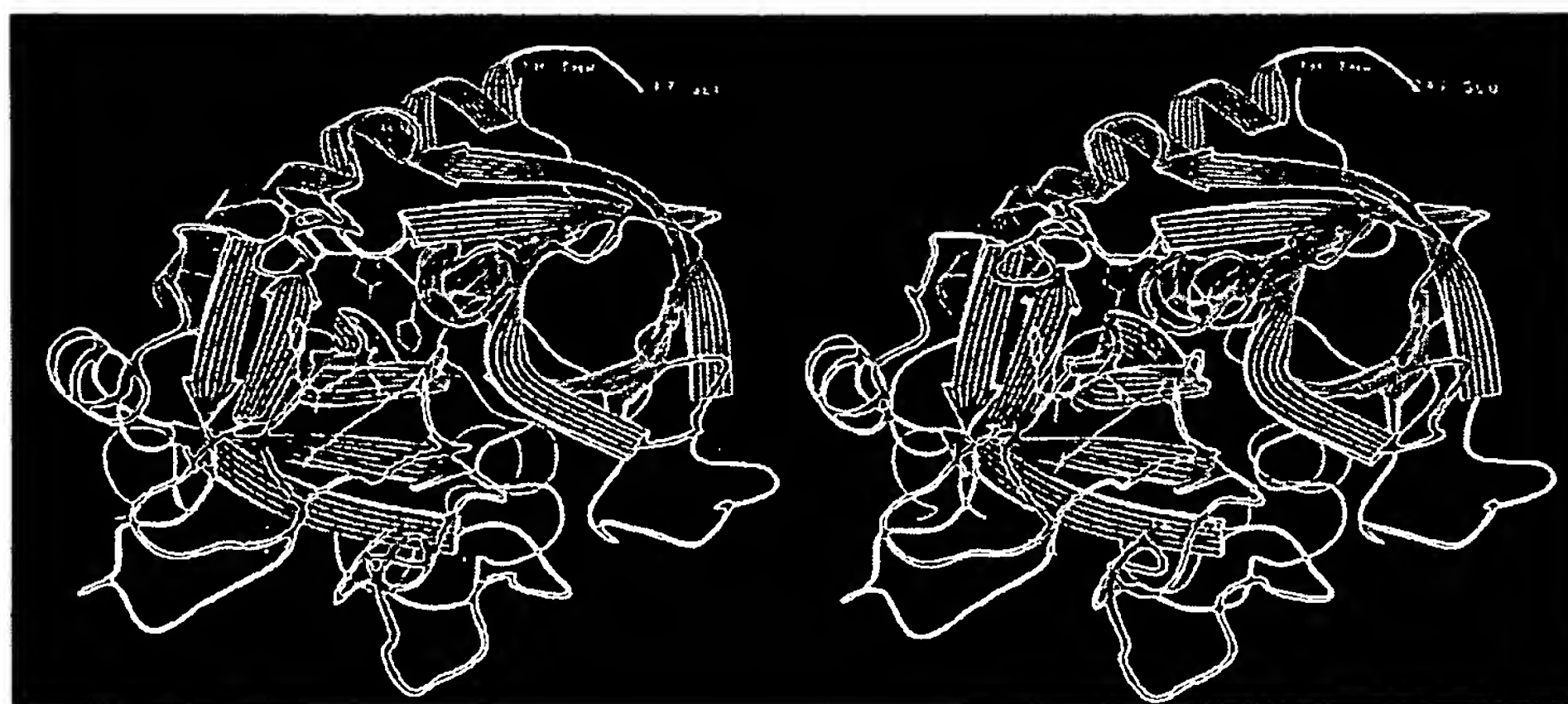


Fig. 2. Ribbon plot of the human  $\alpha$ -thrombin polypeptide chain;  $\beta$ -strands, helices, and turns are represented by twisted arrows, helical ribbons, and ropes. The view is on the active-site cleft made by the B-chain (yellow); the A-chain runs in the back of the molecule (violet). This figure has been produced with a modified version of the program RIBBON kindly provided by John Priestle (Priestle, 1988).



The secondary structure of the A- and the B-chain of the final thrombin model is illustrated schematically in Figure 3. Each barrel-like domain appears as six antiparallel running strands. The size and extent of each  $\beta$ -sheet is remarkably similar to those observed in other trypsin-like serine proteinases. The polypeptide strands involved in barrel formation always fold back on the molecular surface (see Fig. 2); some of these surface-located thrombin loops are particularly extended in comparison to other related serine proteinases. In each barrel, four of the five turns are hairpin loops, all of them of a rather open, nonregular conformation. Some of these open loops contain local classical 1–4 tight turns (Table 1).

The six helical segments occurring in thrombin are listed in Table 2 together with their helix type and length

defined according to different helix criteria. The one-turn helix around His 57, the intermediate helix around Cys 168, and the long carboxy-terminal helix (of a mixed type) are similar to those found in other serine proteinases; the A-chain helix, the short  $3_{10}$ -helix between Tyr 60A and Asp 60E (better described as two interlaced type-III turns; see Table 1), and the regular two-turn  $\alpha$ -helix around Ala 129 are additional segments or parts of insertion loops and seem to be characteristic of thrombin.

#### Comparison with related serine proteinase structures

The  $\alpha$ -carbon structure of human  $\alpha$ -thrombin optimally superimposed on that of bovine chymotrypsinogen is

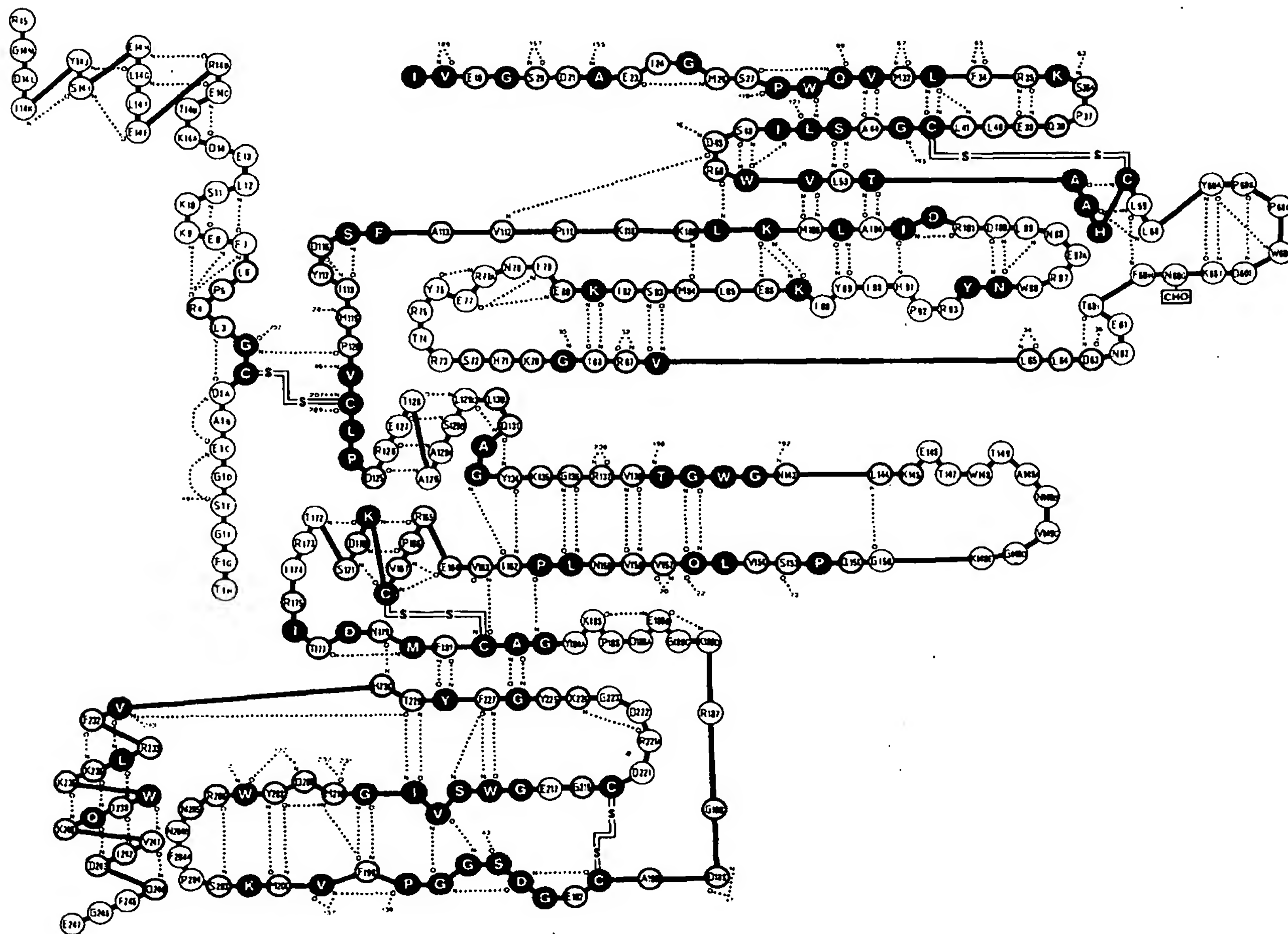


Fig. 3. Secondary structure of human  $\alpha$ -thrombin. Inter-main-chain hydrogen bonds are selected according to criteria (with  $E < -0.7$  kcal/mol) given by Kabsch and Sander (1983). The sequence nomenclature is that derived from equivalency with chymotrypsin(ogen)/trypsin (Bode et al., 1989b) as also used in Table 3. Thrombin residues that are topologically equivalent to chymotrypsin residues are emphasized by thick circles; thrombin residues, which in addition possess identical amino-acids, are denoted by filled circles.



Table 1. Tight 1-4 turns in human  $\alpha$ -thrombin

Residues	Position				Turn type	Hydrogen bond		
	1	2	3	4		1-4 Hydrogen-bond length (Å)	Energy (kcal/mol) <sup>a</sup>	Remarks
1A-3	Asp	Cys	Gly	Leu	II	3.45	-1.3	Gly in position 3
4-7	Arg	Pro	Leu	Phe	I-III	3.01	-1.5	Pro in position 2
8-11	Glu	Lys	Lys	Ser	I	2.89	-1.9	
14-14C	Asp	Lys	Thr	Glu	I	3.21	-1.6	
14G-14J	Leu	Glu	Ser	Tyr	III	2.80	-2.1	
23-26	Glu	Ile	Gly	Met	II	3.40	-1.3	Gly in position 3
48-51	Ser	Asp	Arg	Trp	I	3.46	-0.7	Very weak hydrogen bond
60A-60D	Tyr	Pro	Pro	Trp	III	3.20	-1.5	Part of $3_{10}$ -helix
60B-60E	Pro	Pro	Trp	Asp	I	2.74	-2.8	Pro not in position 3
60I-63	Thr	Glu	Asn	Asp	I	2.97	-1.8	
77-79	Glu	Arg	Asn	Ile	I	3.41	-0.8	Very weak hydrogen bond
91-94	His	Pro	Arg	Tyr	I	3.39	-1.0	Pro not in position 3
131-134	Gln	Ala	Gly	Tyr	II	2.84	-2.4	Gly in position 3
164-167	Glu	Arg	Pro	Val	III	3.07	-0.9	
168-171	Cys	Lys	Asp	Ser	I	3.11	-1.6	
177-180	Thr	Asp	Asn	Met	I	3.12	-1.5	
185-186B	Lys	Pro	Asp	Glu	I	3.01	-2.2	Pro not in position 3
191-194	Cys	Glu	Gly	Asp	II	3.34	-1.5	Gly in position 3
221-224	Arg	Asp	Gly	Lys	II	3.18	-1.7	Gly in position 3

<sup>a</sup> According to Kabsch and Sander (1983).

shown in Figure 4. One hundred ninety amino acid residues of the thrombin B-chain are topologically equivalent (within a root-mean-square [r.m.s.] deviation of 0.78 Å) to residues in bovine chymotrypsin (Cohen et al., 1981; Blevins & Tulinsky, 1985; Tsukada & Blow, 1985); in addition, 6 residues of the thrombin A-chain are equivalent to the 6 amino-terminal residues of the chymotrypsin(ogen) propeptide. The 190 equivalent B-chain residues and another 40 residues were assigned the sequence numbers of the topologically equivalent chymotrypsinogen residues (Hartley & Kauffman, 1966; Meloun et al., 1966). In addition to secondary structural elements, Figure 3 defines those human  $\alpha$ -thrombin amino acid resi-

dues topologically equivalent to chymotrypsin residues (see also Kinemage 2); filled circles further represent those 84 with identical amino acids.

The thrombin B-chain has 195  $\alpha$ -carbon atoms equivalent to the bovine trypsin structure (see Fig. 5) with an r.m.s. deviation of 0.8 Å (Bode & Schwager, 1975a); at a few sites the thrombin structure is closer to that of trypsin than to that of chymotrypsin, i.e., (1) thrombin segment 70-80, which resembles the calcium-binding loop of trypsin (Bode & Schwager, 1975a,b), (2) Gly 184-Tyr 184A, and (3) thrombin loop segment 217-224.

Here, the generally used chymotrypsinogen numbering of bovine trypsin (Hartley & Shotton, 1971) was used.

Table 2. Helical segments in human  $\alpha$ -thrombin

Helix type	Helix extent (number of residues) according to Richardson and Richardson (1990)		
	Hydrogen bonding	Conformation	$\alpha$ -Carbon position
3.6 <sub>13</sub>	Glu 14C to Ser 14I (7)	Glu 14C to Ser 14I (7)	Thr 14B to Tyr 14J (9)
3 <sub>10</sub>	Ala 55 to Leu 58 (5)	Ala 56 to Leu 59 (5)	Ala 56 to Tyr 60A (6)
3 <sub>10</sub>	Tyr 60A to Asp 60E (5)	Pro 60B to Pro 60C (2)	Tyr 60A to Trp 60D (4)
3.6 <sub>13</sub>	Asp 125 to Leu 129C (8)	Arg 126 to Ser 129B (6)	Asp 125 to Glu 131 (10)
3.6 <sub>13</sub>	Glu 164 to Asp 170 (7)	Arg 165 to Lys 169 (5)	Glu 164 to Ser 171 (8)
3 <sub>10</sub>	Cys 168 to Thr 172 (5)	Lys 169 to Ser 171 (3)	Lys 169 to Thr 172 (4)
3 <sub>10</sub>	Val 231 to Lys 235 (5)	Val 231 to Trp 237 (7)	His 230 to Leu 234 (5)
3.6 <sub>13</sub>	Leu 234 to Asp 243 (10)	Ile 238 to Ile 242 (5)	Leu 234 to Ile 242 (9)
3 <sub>10</sub>	Val 241 to Glu 244 (4)	Val 241 to Glu 244 (4)	Val 241 to Glu 244 (4)

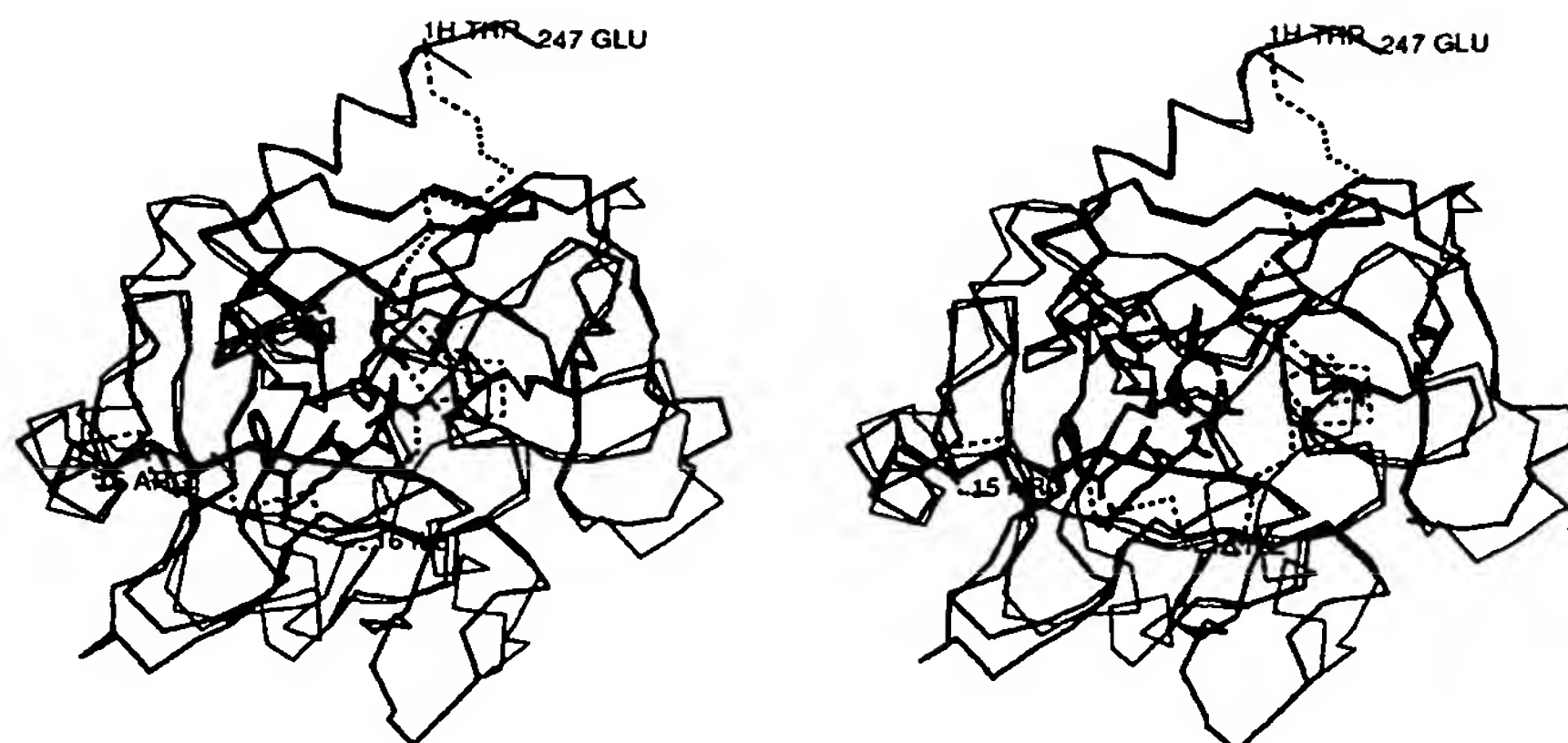


Fig. 4.  $\alpha$ -carbon structure of the A-chain (dashed connections) and the B-chain (thick connections) of human  $\alpha$ -thrombin superimposed with molecule 1 of bovine chymotrypsinogen (thin connections [Wang et al., 1985]) after minimizing the r.m.s. deviation of the topologically equivalent  $\alpha$ -carbon atoms. The PPACK molecule is shown with all nonhydrogen atoms, and the thrombin chain termini are labeled.

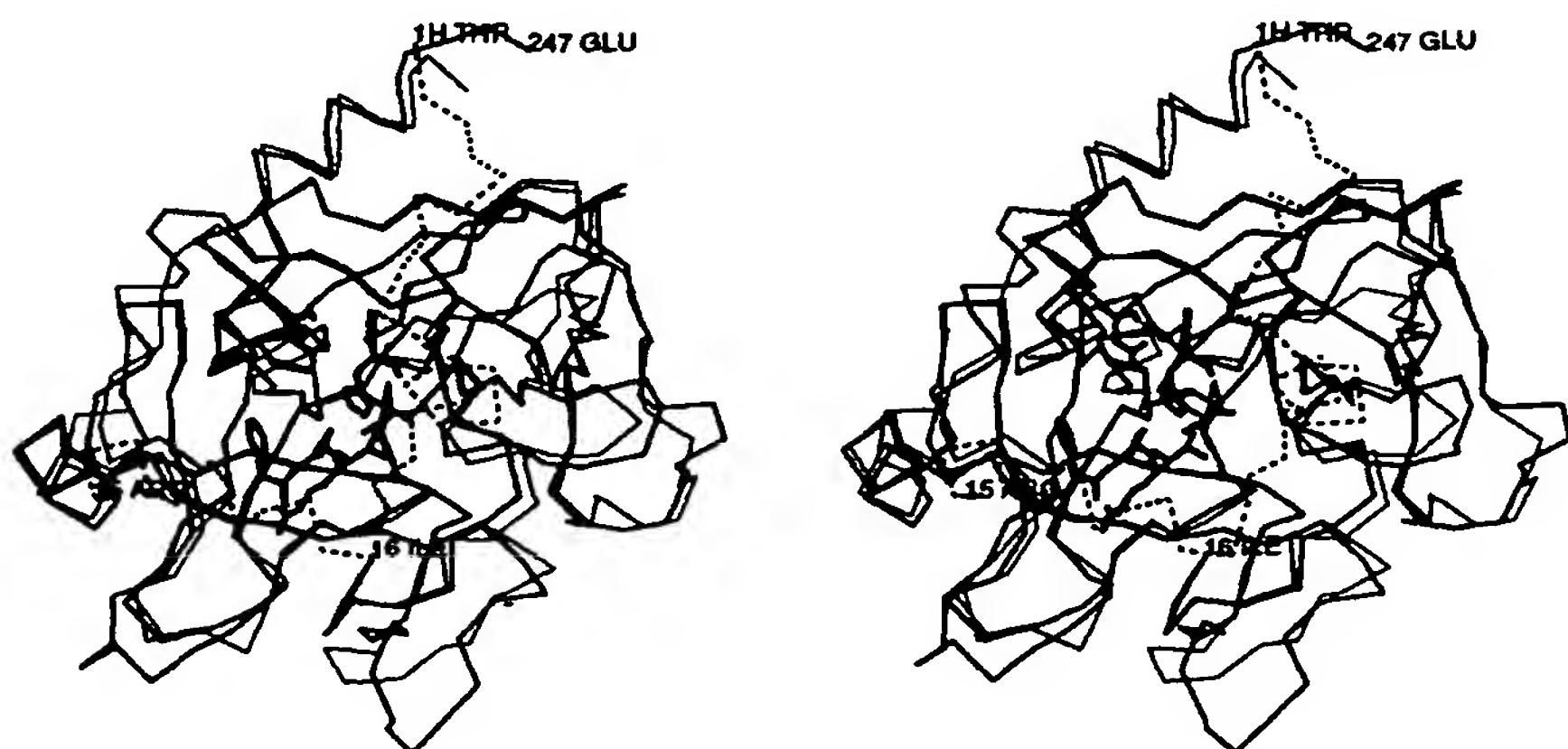


Fig. 5.  $\alpha$ -carbon structure of the A-chain (dashed connections) and the B-chain of human  $\alpha$ -thrombin (thick connections) superimposed with bovine trypsin (Bode et al., 1976). The PPACK molecule is overlaid; standard view as in Figure 1.

The residual insertion residues in the thrombin B-chain were marked by letter suffixes; at a few of the junctions between conserved and variable segments, the positioning of the insertions was necessarily somewhat arbitrary.

For numbering of the A-chain, residue numbers 1 and 15 were assigned to the (chymotrypsin-homologous) cysteine and to the carboxy-terminal arginine residue, respectively. The A-chain residues preceding Cys 1 were designated by a 1 and a letter suffix in alphabetic order in the reverse chain direction; the A-chain residues following Cys 1 were designated by numbers up to 14, followed by 14 and letter suffixes in alphabetic order.

The chymotrypsinogen sequence numbering of human thrombin is shown in Table 3 based on these topological equivalences with chymotrypsin and trypsin (see Bode et al., 1989b). The A-chain runs from Thr 1H to Arg 15; the B-chain starts with Ile 16 and ends up with Glu 247. A nomenclature based on the topological equivalence of the amino acid residues is an enormous advantage, as functional properties already known for the related trypsinlike enzymes can be attributed to distinct sites of the

thrombin structure. Table 3 also contains the sequences of bovine chymotrypsinogen (Hartley & Kauffman, 1966; Meloun et al., 1966), bovine trypsinogen (Walsh & Neurath, 1964; Mikes et al., 1966), and bovine thrombin (Magnusson et al., 1975; MacGillivray & Davie, 1984). All key residues of thrombin are also conserved in the corresponding enzymes from mouse (Friezner-Degen et al., 1990) and rat (Dihanich & Monard, 1990).

According to this topological alignment, 86 (of the equivalent) amino acid residues of the thrombin B-chain are identical with residues in bovine trypsin, and 79 are identical with residues in bovine chymotrypsin (whose propeptide possesses, in addition, two residues identical to the thrombin A-chain). For comparison, the two digestive enzymes have 99 identical residues between them.

Of the other vertebrate serine proteinases of known spatial structure, porcine pancreatic kallikrein (Bode et al., 1983) and porcine pancreatic elastase (Meyer et al., 1988) exhibit a similar topological equivalence to the thrombin B-chain; however, their r.m.s. deviation is slightly larger and the number of identical residues slightly smaller (Ta-

**Table 3.** Sequence alignment of bovine thrombin (BTHR), human thrombin (HTHR), bovine trypsin (BTRY), and bovine chymotrypsin (BCHY) according to topological equivalences; thrombin chymotrypsinogen numbering (THRO), chymotrypsinogen numbering (CHYG)<sup>a</sup>

THRO	1H	1G	1F	1E	1D	1C	1B	1A	1	2	3	4	5	6	7
BTHR	T	F	G	A	G	E	A	D	C	G	L	R	P	L	F
HTHR	T	F	G	S	G	E	A	D	C	G	L	R	P	L	F
BTRY									C	G	V	P	A	I	Q
BCHY									1	2	3	4	5	6	7
CHYG															
THRO	8	9	10	11	12	13	14	14A	14B	14C	14D	14E	14F	14G	14H
BTHR	E	K	K	Q	V	Q	D	Q	T	E	K	E	L	F	E
HTHR	E	K	K	S	L	E	D	K	T	E	R	E	L	L	E
BTRY	P	V	L	S	G	L	S	-	-	-	-	-	-	-	-
BCHY	P	V	L	S	G	L	S	-	-	-	-	-	-	-	-
CHYG	8	9	10	11	12	13	14	-	-	-	-	-	-	-	-
THRO	14I	14J	14K	14L	14M	15									
BTHR	S	Y	I	E	G	R									
HTHR	S	Y	I	D	G	R									
BTRY	V	D	D	D	D	K									
BCHY	-	-	-	-	-	R									
CHYG	-	-	-	-	-	15									
THRO	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
BTHR	I	V	E	G	Q	D	A	E	V	G	L	S	P	W	Q
HTHR	I	V	E	G	S	D	A	E	I	G	M	S	P	W	Q
BTRY	I	V	G	G	Y	T	C	G	A	N	T	V	P	Y	Q
BCHY	I	V	N	G	E	E	A	V	P	G	S	W	P	W	Q
CHYG	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
THRO	31	32	33	34	35	36	36A	37	38	39	40	41	42	43	44
BTHR	V	M	L	F	R	K	S	CP	Q	E	L	L	C	G	A
HTHR	V	M	L	F	R	K	S	CP	Q	E	L	L	C	G	A
BTRY	V	S	L	N	S	-	-	-	G	Y	H	F	C	G	G
BCHY	V	S	L	Q	D	K	-	T	G	F	H	F	C	G	G
CHYG	31	32	33	34	35	36	-	37	38	39	40	41	42	43	44
THRO	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
BTHR	S	L	I	S	D	R	W	V	L	T	A	A	H	C	L
HTHR	S	L	I	S	D	R	W	V	L	T	A	A	H	C	L
BTRY	S	L	I	N	S	Q	W	V	V	S	A	A	H	C	Y
BCHY	S	L	I	N	E	N	W	V	V	T	A	A	H	C	G
CHYG	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
THRO	60	60A	60B	60C	60D	60E	60F	60G	60H	60I	61	62	63	64	65
BTHR	L	Y	P	P	W	D	K	N	H	I	V	D	D	L	L
HTHR	L	Y	P	P	W	D	K	N	H	I	E	N	D	L	L
BTRY	K	-	-	-	-	-	-	-	-	-	-	S	G	I	Q
BCHY	V	-	-	-	-	-	-	-	-	-	T	T	S	D	V
CHYG	-	-	-	-	-	-	-	-	-	-	61	62	63	64	65
THRO	66	67	68	69	70	71	72	73	74	75	76	77	77A	78	79
BTHR	V	R	I	G	K	H	S	R	T	R	Y	E	R	K	V
HTHR	V	R	I	G	K	H	S	R	T	R	Y	E	R	N	I
BTRY	V	R	L	G	E	D	N	I	N	V	V	E	-	G	N
BCHY	V	V	A	G	E	F	D	Q	G	S	S	S	-	E	K
CHYG	66	67	68	69	70	71	72	73	74	75	76	77	77A	78	79
THRO	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94
BTHR	E	K	I	S	M	L	D	K	I	Y	I	H	P	R	Y
HTHR	E	K	I	S	M	L	E	K	I	Y	I	H	P	R	Y
BTRY	E	Q	F	I	S	A	S	K	S	I	V	H	P	S	Y
BCHY	I	Q	K	L	K	I	A	K	V	F	K	N	S	K	Y
CHYG	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94

(continued)

Table 3. Continued

THRO	95	96	97	97A	98	99	100	101	102	103	104	105	106	107	108
BTHR	N	W	K	E	N	L	D	R	D	I	A	L	L	K	L
HTHR	N	W	R	E	N	L	D	R	D	I	A	L	M	K	L
BTRY	N	S	N	-	T	L	N	N	D	I	M	L	I	K	L
BCHY	N	S	L	-	T	I	N	N	D	I	T	L	L	K	L
CHYG	95	96	97	-	98	99	100	101	102	103	104	105	106	107	108
THRO	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123
BTHR	K	R	P	I	E	L	S	D	Y	I	H	P	V	C	L
HTHR	K	K	P	V	A	F	S	D	Y	I	H	P	V	C	L
BTRY	K	S	A	A	S	L	N	S	R	V	A	S	I	S	L
BCHY	S	T	A	A	S	F	S	Q	T	V	S	A	V	C	L
CHYG	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123
THRO	124	125	126	127	128	129	129A	129B	129C	130	131	132	133	134	135
BTHR	P	D	K	Q	T	A	A	K	L	L	H	A	G	F	K
HTHR	P	D	R	E	T	A	A	S	L	L	Q	A	G	Y	K
BTRY	P	T	-	S	C	A	-	-	-	-	S	A	G	T	Q
BCHY	P	S	A	S	D	D	-	-	-	F	A	A	G	T	T
CHYG	124	125	126	127	128	129	-	-	-	130	131	132	133	134	135
THRO	136	137	138	139	140	141	142	143	144	145	146	147	148	149	149A
BTHR	G	R	V	T	G	W	G	N	R	R	E	T	W	T	T
HTHR	G	R	V	T	G	W	G	N	L	K	E	T	W	T	A
BTRY	C	L	I	S	G	W	G	N	T	K	S	S	G	T	-
BCHY	C	V	T	T	G	W	G	L	T	R	Y	T	N	A	-
CHYG	136	137	138	139	140	141	142	143	144	145	146	147	148	149	-
THRO	149B	149C	149D	149E	150	151	152	153	154	155	156	157	158	159	160
BTHR	S	V	A	E	V	Q	P	S	V	L	Q	V	V	N	L
HTHR	N	V	G	K	G	Q	P	S	V	L	Q	V	V	N	L
BTRY	-	-	-	-	S	Y	P	D	V	L	K	C	L	K	A
BCHY	-	-	-	-	N	T	P	D	R	L	Q	Q	A	S	L
CHYG	-	-	-	-	150	151	152	153	154	155	156	157	158	159	160
THRO	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175
BTHR	P	L	V	E	R	P	V	C	K	A	S	T	R	I	R
HTHR	P	I	V	E	R	P	V	C	K	D	S	T	R	I	R
BTRY	P	I	L	S	D	S	S	C	K	S	A	Y	P	G	Q
BCHY	P	L	L	S	N	T	N	C	K	K	Y	W	G	T	K
CHYG	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175
THRO	176	177	178	179	180	181	182	183	184	184A	185	186	186A	186B	186C
BTHR	I	T	D	N	M	F	C	A	G	Y	K	P	G	E	G
HTHR	I	T	D	N	M	F	C	A	G	Y	K	P	D	E	G
BTRY	I	T	S	N	M	F	C	A	G	Y	L	E	-	-	-
BCHY	I	K	D	A	M	I	C	A	G	-	A	S	-	-	-
CHYG	176	177	178	179	180	181	182	183	184	-	185	186	-	-	-
THRO	186D	187	188	189	190	191	192	193	194	195	196	197	198	199	200
BTHR	K	R	G	D	A	C	E	G	D	S	G	G	P	F	V
HTHR	K	R	G	D	A	C	E	G	D	S	G	G	P	F	V
BTRY	G	G	K	D	S	C	Q	G	D	S	G	G	P	V	V
BCHY	-	G	V	S	S	C	M	G	D	S	G	G	P	L	V
CHYG	-	187	188	189	190	191	192	193	194	195	196	197	198	199	200
THRO	201	202	203	204	204A	204B	205	206	207	208	209	210	211	212	213
BTHR	M	K	S	P	Y	N	N	R	W	Y	Q	M	G	I	V
HTHR	M	K	S	P	F	N	N	R	W	Y	Q	M	G	I	V
BTRY	C	S	-	-	-	-	-	-	G	K	L	Q	G	I	V
BCHY	C	K	K	N	-	-	G	A	W	T	L	V	G	I	V
CHYG	201	202	203	204	-	-	205	206	207	208	209	210	211	212	213

(continued)



Table 3. Continued

THRO	214	215	216	217	-	219	220	221	221A	222	223	224	225	226	227
BTHR	S	W	G	E	-	G	C	D	R	D	G	K	Y	G	F
HTHR	S	W	G	E	-	G	C	D	R	D	G	K	Y	G	F
BTRY	S	W	G	S	-	G	C	A	Q	K	N	K	P	G	V
BCHY	S	W	G	S	S	T	C	S	-	T	S	T	P	G	V
CHYG	214	215	216	217	218	219	220	221	-	222	223	224	225	226	227

THRO	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242
BTHR	Y	T	H	V	F	R	L	K	K	W	I	Q	K	V	I
HTHR	Y	T	H	V	F	R	L	K	K	W	I	Q	K	V	I
BTRY	Y	T	K	V	C	N	Y	V	S	W	I	K	Q	T	I
BCHY	Y	A	R	V	T	A	L	V	N	W	V	Q	Q	T	L
CHYG	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242

THRO	243	244	245	246	247
BTHR	D	R	L	G	S
HTHR	D	Q	F	G	E
BTRY	A	S	N	-	-
BCHY	A	A	N	-	-
CHYG	243	244	245	246	247

<sup>a</sup> Residues topologically equivalent with human thrombin are boxed.

ble 4). The corresponding numbers (Table 4) for rat mast cell proteinase II (Remington et al., 1988) and human leukocyte elastase (Bode et al., 1986), two cellular serine proteinases, indicate a more distant structural relationship to thrombin.

This ranking order due to topological equivalence agrees approximately with the order obtained from a sequence alignment of these serine proteinases with the

thrombin B-chain (Table 4). Table 4 shows furthermore the alignment score and the degree of similarity to the thrombin B-chain for the reactive domains of several coagulation factors based on sequence alignment alone. According to such data, the human thrombin B-chain is most similar to the catalytic domains of protein C and Factor Xa of human origin; its similarity with coagulation Factors IXa and XIa is only slightly lower, in agree-

Table 4. Topological equivalence and sequence identity of human thrombin B-chain with the reactive domains of vertebrate serine proteinases. Sequence alignment scores for other coagulation factors<sup>a</sup>

	Topologically equivalent residues		Identical residues (from sequence alignment)		
	Number	r.m.s. deviation (Å)	Number	% Total (enzyme)	Alignment score (standard deviations)
BTRY	195	0.82	88	39.5	27
BCHY	190	0.78	85	37.0	24
PPKK	193	0.92	70	30.2	19
PPE	192	1.02	62	25.8	12
RMCP	180	0.89	56	25.0	13
HLE	176	0.90	61	28.0	12
(BTRY-BCHY)	190	0.76	96	43.0	35
Protein C			100	40.0	41
FXa (reactive domain)			97	38.2	40
FIXa (reactive domain)			87	37.0	37
FXIa (reactive domain)			88	37.0	34

<sup>a</sup> BTRY, bovine trypsin (Bode & Schwager, 1975a); BCHY, bovine chymotrypsin A (Cohen et al., 1981); PPKK, porcine pancreatic kallikrein (Bode et al., 1983); RMCP, rat mast cell proteinase II (Remington et al., 1988); PPE, porcine pancreatic elastase (Meyer et al., 1988); HLE, human leukocyte elastase (Bode et al., 1986).

ment with previous investigations of others (see, e.g., Patthy, 1985).

The polypeptide alignment (Table 3) obtained for thrombin and chymotrypsin(ogen) due to topological equivalence is in good agreement with the alignment scheme proposed by Hartley and Shotton (1971). The topologically equivalent regions of the thrombin B-chain (Table 3) contain the (structurally) conserved regions previously identified by Greer (1981) and Furie et al. (1982) on the basis of amino acid sequence homology or from comparison of the three-dimensional models of serine proteinases, respectively. The sites and numbers of insertions and deletions predicted by Bing et al. (1986) on the basis of a chymotrypsin-trypsin hybrid model are largely in accord with our experimental results. Their proposed alignment turned out to be incorrect only around Glu 80; as shown below, a single insertion at Arg 77A is found compared with the predicted deletion and insertion at positions 76 and 84, respectively. Some spatial properties of thrombin such as the positively charged anion-binding exosite and the clustering of aromatic residues near the active site were correctly predicted (Bing et al., 1986). It had been suggested that one or both large insertion loops might fold (partially) over the entrance to the catalytic center, in this way blocking the approach of large macromolecular substrates or inhibitors (Sugawara et al., 1986). Not foreseen was the role of Ile 174 (which is not part of an insertion segment, see below) or some other structural details largely responsible for thrombin specificity.

#### Comparison with related thrombin structures

The availability of different crystal structures of thrombin from different species allows us to recognize intrinsic structural properties and to separate them from crystal-packing effects. Comparison of the  $\alpha$ -carbon tracing of PPACK-thrombin with that of the hirudin-thrombin complex (Rydel et al., 1991; see Fig. 6) shows both thrombin models to be to a large extent identical; two-thirds of all  $\alpha$ -carbon atoms (i.e., 200 largely internal res-

idues) differ by only 0.26 Å (r.m.s. deviation). There are, however, some very significant displacements and conformational differences, in particular in the amino-terminal segment (up to Ala 1B) and in the carboxy-terminal segment (from Ile 14K onward) of the A-chain, in the insertion loop around Trp 148, and at the carboxy-terminus of the B-chain from Gln 244 onward.

The displacements of the residues around Trp 148 are the result of a large conformational change of the corresponding loop, mainly through a rigid-body rotation of the loop domain around Glu 146 C $\alpha$ -C and Gly 150 N-C $\alpha$  (see Fig. 17 and below). The central residues of this loop are displaced by up to 10 Å (Rydel et al., 1991). Without rearrangements, the hirudin chain as it is observed in the hirudin-thrombin complex would collide with the Trp 148 loop in its PPACK-thrombin conformation (in particular, the side chains of Trp 148 and Thr 4IH [see Fig. 17]). It should be mentioned that this loop segment in both thrombin structures exhibits enhanced, but by no means extreme, B-values (see Fig. 28).

Significant deviations (up to almost 1.0 Å) are furthermore observed at the central residues of the insertion loop around Trp 60D (which juxtaposes the Trp 148 loop across the active-site cleft). These displacements result, however, from expansion of the S2 cavity (see below) enclosed by this exposed loop rather than from conformational changes. Again, the B-values of this loop are not exceptionally high in either structure.

Further significant, but more local, positional changes between PPACK-thrombin and hirudin-thrombin are observed at Glu 97A (which forms the ceiling of the hydrophobic pocket) and at Arg 77A (in the fibrin[ogen] secondary binding exosite); both residues are parts of quite exposed loops in the substrate-binding area and are rather flexible; thus their conformation might easily be affected by the differently bound inhibitors.

Comparison of the human PPACK-thrombin model with the two independent molecules (Martin, Kunjumen, Kumar, Bode, Huber, & Edwards, unpubl.) in the bovine thrombin structure (with 87% identical residues

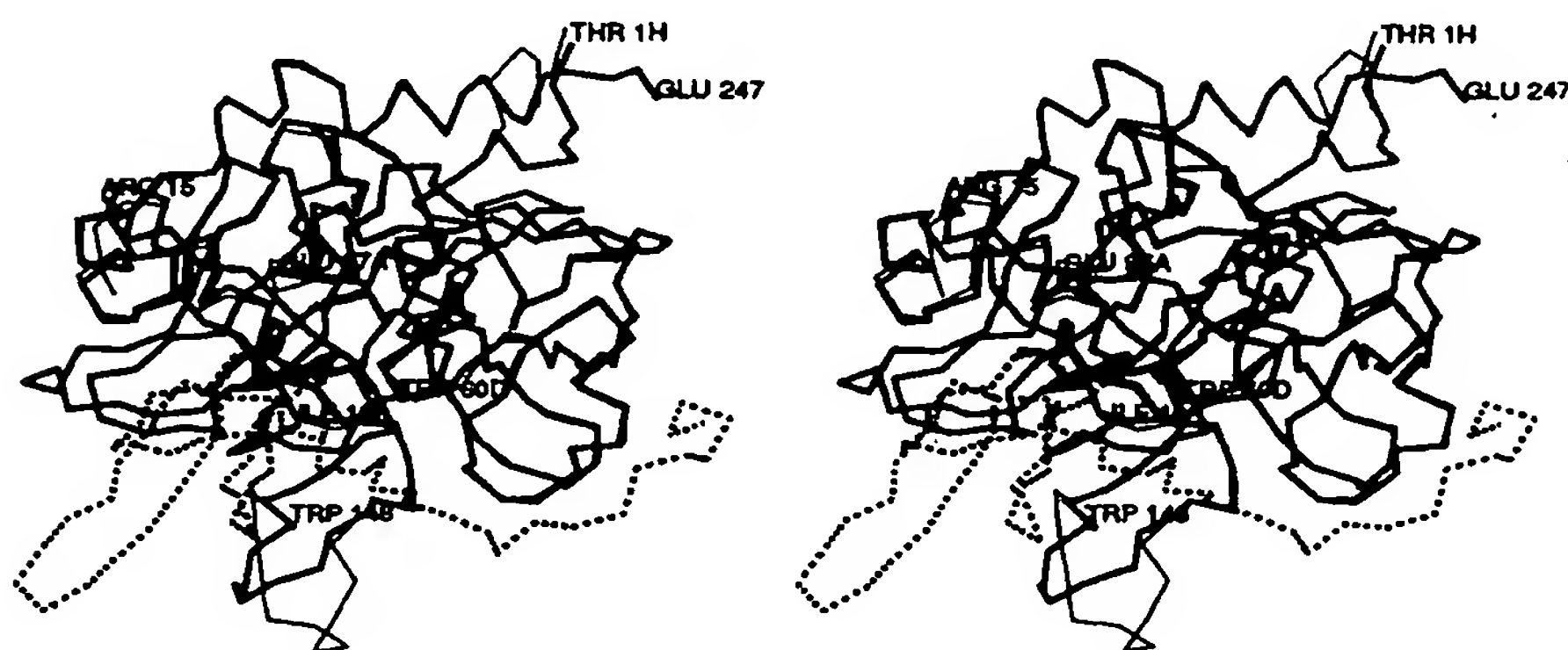


Fig. 6.  $\alpha$ -carbon structure of PPACK-human  $\alpha$ -thrombin (thick connections) optimally superimposed with the thrombin (thin connections) and the hirudin component (dashed connections) of the thrombin-hirudin complex (Rydel et al., 1991). Standard view as in Figure 1.

[compare with Table 3]) yields a similar pattern of polypeptide chain displacements: the free amino- and the free carboxy-terminal end of the A-chain, the carboxy-terminus of the B-chain, and the Trp 148 loop (which is only weakly defined presumably due to partial cleavage) show very large displacements, mainly as the result of different crystal contacts. Omission of these segments (i.e., of 21 amino acid residues) yields an r.m.s. deviation of only 0.45 Å for  $\alpha$ -carbon atoms. It is noteworthy that the 60-insertion loop is similarly folded in the two bovine thrombin molecules as in PPACK-thrombin.

#### *Charged groups and salt bridges*

Human thrombin has a particularly high proportion of polar amino acids. Only about 45% of its total residues are nonpolar (compared with, for example, 51% in bovine chymotrypsin and 63% in the much more "sticky" human leukocyte elastase); more than 70% of all residues on the thrombin surface are hydrophilic in nature. Moreover, human  $\alpha$ -thrombin possesses an exceptionally large number of charged residues compared with the pancreatic trypsinlike enzymes. The A- and the B-chain contain 9 and 31 acidic residues and 6 and 37 basic residues (and an additional 4 histidines), respectively (Fig. 7). In contrast, bovine trypsinogen and bovine chymotrypsinogen contain only 12 or 14 acidic, and 17 or 18 basic residues. In addition, human  $\alpha$ -thrombin has two terminal amino, two terminal carboxy, and two neuraminic acid groups, the latter terminating the single Asn 60G-linked sugar chain of thrombin (Nilsson et al., 1983).

The calculated charges and electrostatic energies of all (partially) charged protein groups in the electrostatic multipole field created by the whole thrombin molecule are listed in Table 5. About three-quarters of all charged groups exhibit negative electrostatic interaction energies, i.e., confer overall stability to the thrombin molecule, whereas one-quarter of them have a destabilizing influ-

ence on the molecular structure. At pH 7, His 57 should carry an almost full positive charge (due to the close proximity of Asp 102); His 119 should be weakly charged at this pH value, whereas the residual two histidine residues, His 71 and His 91, should virtually be noncharged.

The almost balanced number of acidic and basic residues of human thrombin suggests that its overall net charge (which we calculate to be +2.3 at pH 7, if the two neuraminic acid groups are included) is close to zero under physiological buffer conditions. Considering intrinsic pK values for the ionizable groups of  $\alpha$ -thrombin, an overall isoelectric point of 7.8 can be calculated; allowing for the spatial arrangement of the charged groups and their environment (Karshikov et al., 1989) and assuming two neuraminic acid residues located 20 Å apart from the thrombin surface, a higher value of 8.4 is obtained. These calculated isoelectric points are slightly higher than the experimental values of 7.3 and 7.6 obtained by isoelectric focusing (Fenton et al., 1977); however, higher values around 9 have also been reported for human thrombin (Berg et al., 1979; Heuck et al., 1985).

Seventy-nine of the 89 fully charged groups (at pH 7) are at least partially exposed to bulk water (Table 5); they represent 42% of all surface-located residues of the thrombin molecule. This large number of surface charges presumably accounts for the marked tendency of human thrombin to associate at low ionic strength and for its strongly salt-dependent solubility (Landis et al., 1981). The significantly better solubility of the almost equally charged bovine species indicates that some individual surface charges might have a particular impact on the association behavior, however.

Ten charged protein groups of human  $\alpha$ -thrombin are completely buried and inaccessible to bulk water (Table 5). All of these buried charges and many of the surface-located charges are arranged in pairs and clusters of oppositely charged groups. Thirty-two salt bridges (of stabilizing energies greater than or equal to  $1RT$ , i.e.,

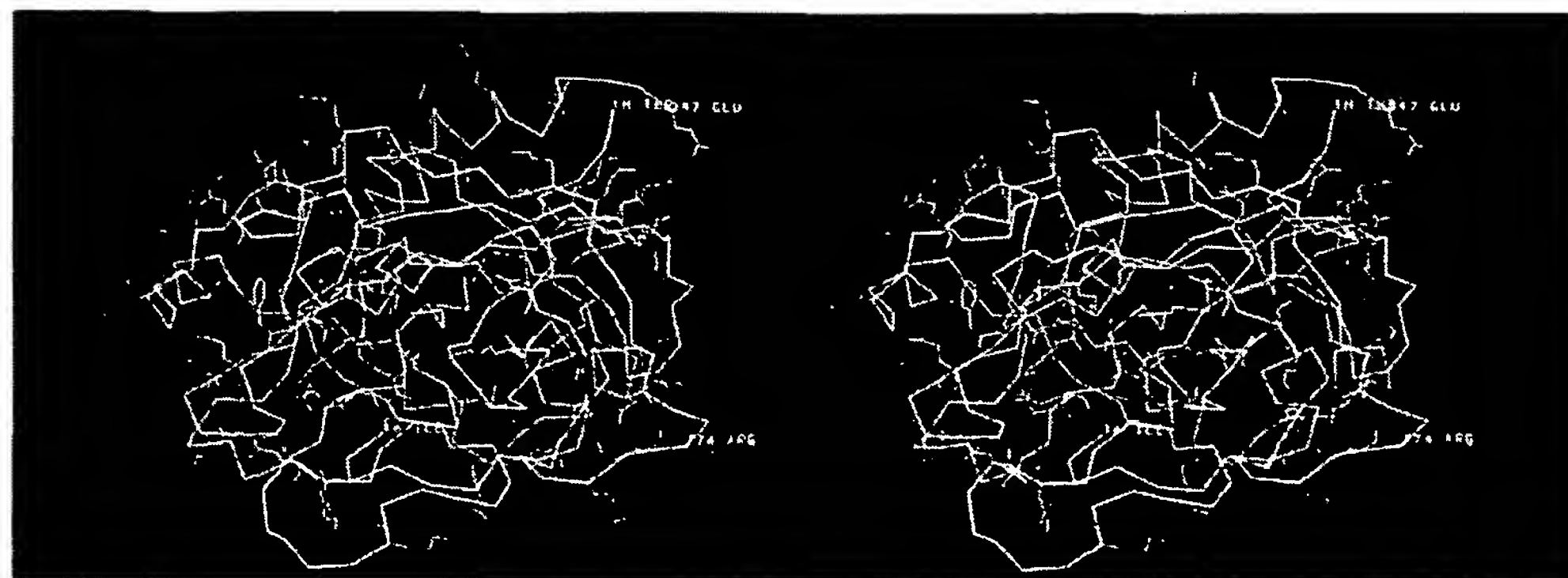


Fig. 7. Main-chain structure of human thrombin (yellow), with all positively (blue) and negatively (red) charged residues added with full structure. Standard view as in Figure 1.

Table 5. Calculated parameters of all charged groups of human  $\alpha$ -thrombin at pH 7, ionic strength 0.125

Residue	Charge	Electrostatic energy (kcal/mol)	Solvent accessibility of charged group	Residue	Charge	Electrostatic energy (kcal/mol)	Solvent accessibility of charged group
Glu 80	-1.0	-4.6	0.00	Asp 21	-1.0	-0.6	0.65
Lys 202	1.0	-4.4	0.13	Asp 125	-1.0	-0.5	0.35
Asp 14	-1.0	-3.8	0.00	Glu 146	-1.0	-0.5	0.06
NTR <sup>a</sup> 16	1.0	-3.6	0.00	Lys 235	1.0	-0.5	0.36
Arg 187	1.0	-3.6	0.33	Asp 116	-1.0	-0.5	0.71
Arg 137	1.0	-3.2	0.00	Lys 10	1.0	-0.4	0.76
Asp 100	-1.0	-3.2	0.06	Glu 1C	-1.0	-0.3	0.27
Arg 4	1.0	-3.1	0.00	Lys 60F	1.0	-0.3	0.28
Lys 185	1.0	-2.8	0.35	His 119	0.3	-0.2	0.05
Glu 77	-1.0	-2.7	0.12	Lys 145	1.0	-0.2	0.28
Arg 50	1.0	-2.6	0.34	Glu 13	-1.0	-0.1	0.54
Asp 194	-1.0	-2.4	0.00	Arg 173	1.0	-0.1	0.98
Glu 8	-1.0	-2.3	0.00	CTR <sup>b</sup> 247	-1.0	-0.1	0.83
Lys 70	1.0	-2.2	0.00	Asp 60E	-1.0	-0.1	0.79
Asp 1A	-1.0	-2.1	0.36	Asp 170	-1.0	-0.1	0.70
Glu 14C	-1.0	-2.0	0.07	Glu 127	-1.0	-0.1	0.76
Lys 224	1.0	-2.0	0.58	Arg 67	1.0	-0.0	0.21
Lys 135	1.0	-1.9	0.24	His 71	0.1	-0.0	0.26
Arg 221A	1.0	-1.7	0.64	His 91	0.0	0.0	0.18
Glu 14E	-1.0	-1.7	0.24	His 230	0.0	0.0	0.40
Lys 9	1.0	-1.7	0.32	Arg 97	1.0	0.0	0.79
Glu 39	-1.0	-1.7	0.49	Arg 75	1.0	0.0	0.91
Asp 178	-1.0	-1.6	0.68	Glu 186B	-1.0	0.0	0.44
Asp 102	-1.0	-1.6	0.00	Glu 14H	-1.0	0.1	0.73
Glu 61	-1.0	-1.4	0.29	Arg 77A	1.0	0.1	0.73
Lys 14A	1.0	-1.4	0.01	CTR 15	-1.0	0.2	0.81
Glu 86	-1.0	-1.2	0.57	Lys 149E	1.0	0.2	0.51
His 57	0.8	-1.2	0.24	Lys 110	1.0	0.2	0.96
Glu 97A	-1.0	-1.1	0.49	Arg 165	1.0	0.2	0.25
Arg 35	1.0	-1.1	0.47	Asp 14L	-1.0	0.3	0.82
Asp 49	-1.0	-1.0	0.19	Lys 81	1.0	0.3	0.56
Lys 107	1.0	-1.0	0.45	Lys 169	1.0	0.3	0.17
Glu 217	-1.0	-1.0	0.33	Arg 233	1.0	0.3	0.54
Glu 164	-1.0	-0.9	0.41	Lys 36	1.0	0.3	0.36
Lys 87	1.0	-0.9	0.59	Lys 240	1.0	0.3	0.89
Arg 206	1.0	-0.9	0.30	Asp 186A	-1.0	0.3	0.71
Arg 14D	1.0	-0.9	0.87	Glu 192	-1.0	0.4	0.53
NTR 1H	1.0	-0.9	0.36	Lys 109	1.0	0.4	0.53
Glu 247	-1.0	-0.8	0.22	Glu 18	-1.0	0.4	0.89
Lys 186D	-1.0	-0.8	0.51	Arg 101	1.0	0.4	0.19
Asp 63	-1.0	-0.8	0.46	Arg 73	1.0	0.6	0.36
Arg 175	1.0	-0.8	0.35	Arg 126	1.0	0.6	0.72
Asp 243	-1.0	-0.7	0.47	Lys 236	1.0	0.7	0.52
Asp 222	-1.0	-0.7	0.53	Asp 221	-1.0	0.7	0.17
Glu 23	-1.0	-0.6	0.29	Asp 189	-1.0	1.1	0.09
Arg 15	1.0	-0.6	0.66	Arg 93	1.0	1.4	0.72

<sup>a</sup> Amino-terminus.<sup>b</sup> Carboxy-terminus.

-0.6 kcal/mol) formed between oppositely charged protein groups are given in Table 6. They contribute to the stabilization of both  $\beta$ -barrels and of some of the open turns. Several of the salt bridges are further interconnected giving rise to ionic clusters of up to four protein groups of either charge. The electrostatic energies of the resulting clusters are, of course, slightly less than the

arithmetic sum of the contributing ion pairs, due to partial repulsion of identically charged groups.

These electrostatic interactions play particularly important roles in the intrachain stabilization of the A-chain and in the coherence between the A- and the B-chain (see Figs. 9, 10; Table 6). Twenty salt bridges (some of which are further organized in five larger ion pair clusters) are



**Table 6.** Ion pairs and clusters stabilizing the tertiary structure of human  $\alpha$ -thrombin

Ion clusters and salt bridges	Electrostatic energy of bridges and clusters (kcal/mol)	Distance <sup>a</sup> (Å)	Solvent accessibility of charged groups involved
<b>Ion pairs stabilizing A-B interactions</b>			
1. Asp 1A-Arg 206	-0.7	6.0/4.2	0.33
2. Glu 8	-1.9	3.8/2.8	0.06
Glu 14C } Lys 202	-3.0		
3. Asp 14-Arg 137	-3.4	2.8/2.8	0.00
4. Glu 14E	-1.6	3.8/2.8	0.33
Lys 135	-1.1		
Lys 186D	-1.1	4.0/3.6	
5. Lys 14A-Glu 23	-1.0	4.5/3.6	0.15
<b>Ion pairs stabilizing A-chain</b>			
1. Asp 1A-Lys 9	-1.8	3.0/2.7	0.34
2. Lys 14A	-0.8	6.0/5.3	0.0
Arg 4 } Asp 14	-2.1		
Glu 8	-1.8	3.7/2.9	
3. Glu 13-Arg 14D	-0.6	4.4/3.3	0.71
<b>Ion pairs stabilizing B-chain</b>			
1. NTR <sup>b</sup> 16-Asp 194	N <sup>c</sup> -3.0	3.0/2.7	0.00
2. Arg 35-Glu 39	T -1.3	3.8/2.7	0.33
3. Arg 50	Asp 49 T -1.5	4.0/3.2	0.25
Glu 247 C -1.5	-1.5		
4. His 57-Asp 102	A -1.3	3.9/2.6	0.12
5. Glu 61-Lys 87	S -1.2	4.1/3.4	0.44
6. Arg 67	-1.9	3.8/3.1	0.08
Lys 70 } Glu 80 T -2.5	-2.5		
Glu 77	-2.1	3.6/2.5	
7. Glu 86-Lys 107	S -0.9	4.7/4.5	0.51
8. Glu 97A	-0.9	4.1/2.7	0.27
Arg 175 D -0.8	-0.8		
Asp 100	-1.8	5.4/5.2	
Arg 101	-1.8	3.9/3.1	
9. Glu 146-Arg 221A	S -1.4	3.9/2.8	0.35
10. Glu 164-Lys 185	S -1.5	3.5/2.7	0.38
11. Arg 165	-0.6	5.1/3.3	0.49
Arg 233 } Asp 178 S -0.7	-0.7		
12. Arg 187	Asp 222 S -1.7	3.0/2.8	0.34
Asp 221 T -1.2	-1.2		
13. Glu 217-Lys 224	S -1.4	3.7/3.4	0.46

<sup>a</sup> Distances between charged group centers (see Materials and methods)/nearest charged atoms.

<sup>b</sup> Amino-terminus.

<sup>c</sup> Fixation of amino-terminus (N) or carboxy-terminus (C); stabilization of turns (T),  $\beta$ -sheets (S); cross-connection of domains (D) or active-site residues (A).

made by B-chain groups alone. In contrast, the comparatively short A-chain is cross-linked by six salt bridges grouped in three separate clusters; another six salt bridges cross-connect the A- with the B-chain (Fig. 10). Altogether, seven of these salt bridges (two within the A-chain, three between A and B, and two within the B-chain) are completely buried in the protein matrix; in each of these buried ion bridges/clusters, at least one of the charged groups is in direct contact with an adjacent internal water molecule (see Table 9).

Particularly extended ion clusters stretch around Asp 14 (where they confer stabilization to the A-chain and tighten the A-B interaction, see below) and Lys 70 (where they stabilize the structure of loop 70-80 and thus maintain the integrity of the fibrin[ogen] secondary binding exosite; see below). Similar to the other activatable trypsinlike serine proteinases, thrombin possesses a buried salt bridge of large electrostatic interaction energy between the ammonium group of Ile 16 (the amino-terminal residue of the B-chain) and the side-chain carboxyl group of Asp 194 (see below and Kinemage 1); the integrity of this salt bridge has been shown to be important in maintaining the characteristic active conformation (see below and Bode, 1979).

The overall distribution of positively and negatively charged residues in the thrombin molecule and along its molecular surface is not uniform (Fig. 7); rather the positive charges are somewhat more concentrated at two poles, and the negatively charged residues arranged in a ring, leading to a sandwichlike electric-field distribution (Fig. 8). In other words, several charges are not properly balanced by oppositely charged protein neighbors, but cluster in positively and negatively charged surface patches (Table 7). These clusters give rise to quite high electrostatic field strengths (both positive and negative) outside the thrombin surface (Fig. 8). The charged groups of residues located in the center of these patches can exhibit considerable positive electrostatic interaction energies with the electric multipole of the molecule (such as, e.g., Arg 93 or Asp 189 [see Tables 5, 7]) and thus have a destabilizing effect on the molecule; several of the charged residues located more toward the periphery of these patches are, on the other hand, simultaneously engaged in salt bridges and can therefore contribute to the stability of the whole molecule (as, e.g., Glu 39 of the negative patch n1a [Table 7]). This charge clustering is presumably the main reason for the strong electrostatic interaction of thrombin with anionic structures observed even at physiological pH values, where its overall charge almost cancels out.

Three such surface patches (patches p1, p2, n1, and subgroups in Table 7) deserve particular attention. The strongest positive electrostatic field (patch p2) is observed for a groove (on top of the thrombin molecule in Fig. 8) surrounded by the exposed side chains of Arg 126, Lys 236, Lys 240, and Arg 93, and of the electrostatically

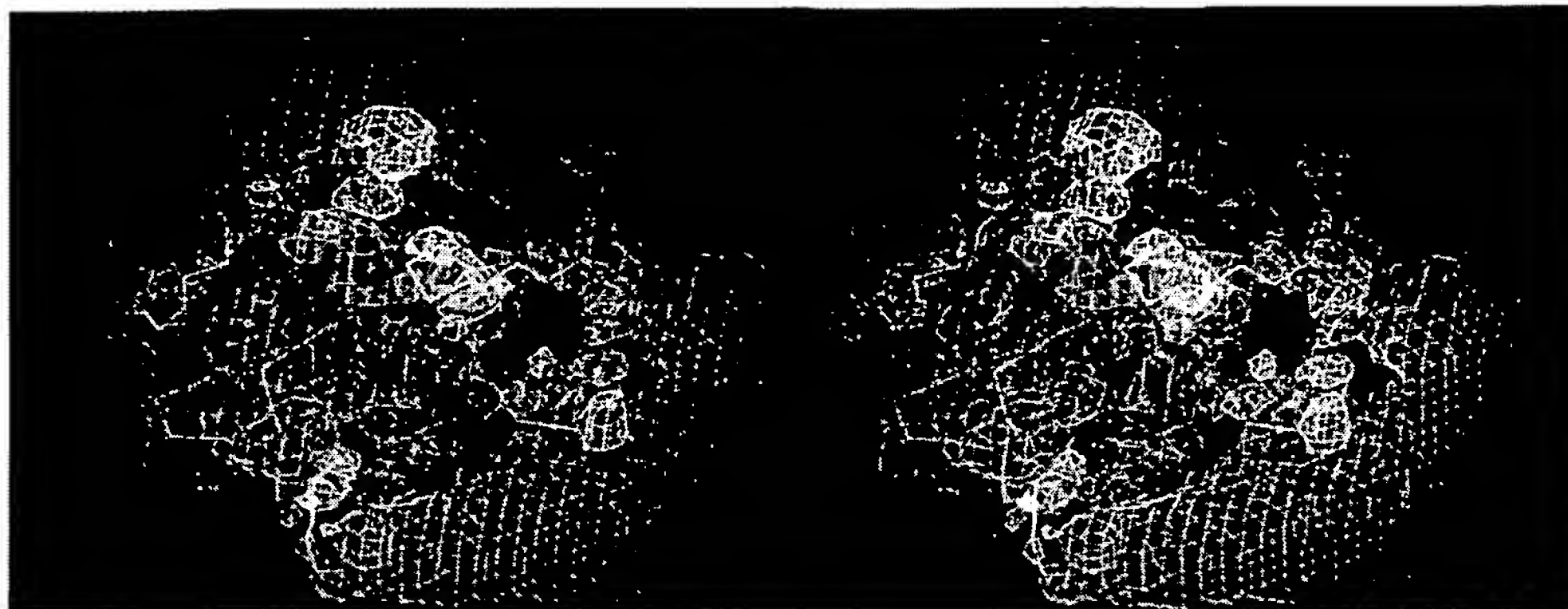


Fig. 8. Main chain structure of human thrombin overlaid with a positive (+3 kcal/mol) and a negative (−5 kcal/mol) electrostatic potential surface. Standard view as in Figure 1.

more compensated residues Lys 235, Arg 101, and Arg 233 in the more distant neighborhood (see Fig. 18). This site presumably represents the heparin-binding site of thrombin and will be discussed further below.

A second surface region of high positive charge density (patch p1) extends along the convex thrombin surface between Lys 149E and Lys 110 (on the right-hand side of the active-site cleft [Fig. 8]). This positively charged patch carries nine basic amino acid side chains (Lys 149E, Arg 67, Arg 73, Arg 75, Arg 77A, Lys 81, Lys 36, Lys 109, and Lys 110) whose charges are only partially compensated by neighboring carboxylate groups. Also this fibrin(ogen) secondary binding exosite will be shown in more detail below.

One larger surface patch of negatively charged residues (Figs. 7, 8) extends around the active site of thrombin (patch n1). It consists of seven aspartic and glutamic acid residues arranged within (Asp 189) and around the specificity pocket: Glu 192 (without protein counter-charge); Asp 102 (active-site residue, hydrogen bonded with His 57); Asp 194 (involved in a relatively strong internal salt bridge with the ammonium group of Ile 16); and Glu 217, Glu 146, and Glu 97A, each involved in salt bridges with positively charged residues in the neighborhood. The many negative charges certainly assist in orienting approaching substrates and affect binding; the electrostatic nature of the catalytic residues located at the edge of this patch is governed more by their adjacent neighbors, however (see below).

#### *Solvent molecules and selected residues*

Thrombin possesses four disulfide bridges (see Kinemage 2). The first (Cys 1–Cys 122) links the A- and B-chains, equivalent topologically to that in chymotrypsin(ogen) (Table 8). Disulfide bridge Cys 42–Cys 58 has a similar conformation to that observed in all other vertebrate ser-

ine proteinases of known tertiary structure. Cys 168–Cys 182 exhibits the same hand in thrombin as in chymotrypsin but opposite to that in trypsin and kallikrein (Bode *et al.*, 1983); however, this bridge differs in detail from all other related proteinases in spite of a quite similar environment. The fourth disulfide bridge, Cys 191–Cys 220, is again of a similar conformation in thrombin as in other related proteinases.

The 409 (423; see Table 15) localized solvent molecules in the PPACK–thrombin structure account for almost 30% of all solvent (water) molecules in the crystals. Their B-values range from 3 Å<sup>2</sup> to 108 Å<sup>2</sup> (3 Å<sup>2</sup> to 88 Å<sup>2</sup>), the average B-value is 50 (45) Å<sup>2</sup>. Thirty-four of these solvent molecules (Table 9) are located within the protein domain (and are in direct contact with at most one bulk solvent molecule); they are considered to be an integral part of the thrombin structure. Seventeen of them are found to be topologically equivalent (i.e., with similar protein environment and deviations up to 1.2 Å after optimal superposition of both protein structures) to waters found in benzamidine-free bovine trypsin (DEBA [Bode *et al.*, 1976; see also Table 3 in Bode & Schwager, 1975a]).

Several of these internal solvents are grouped together forming linear water clusters. A particularly large cluster is made by eight adjacent internal solvent molecules (solvent molecules 315, 316, 320, 323, 342, 343, 345, and 414 [Table 9]) encircled by the thrombin insertion loop Tyr 184A–Gly 188 and segment Asp 221–Tyr 225 behind the side chain of Asp 189; Sol 323 mediates the connection to the bulk water (see Fig. 19). Another large internal solvent cluster comprises five water molecules, extending around Sol 360 close to the 70–80 loop. Sol 305 in the back of the specificity pocket (see Fig. 22) is fully separated from bulk solvent molecules in PPACK–thrombin by the Arg 31 side chain; as with the equivalent Sol 416 in trypsin (Bode & Schwager, 1975a), it plays an impor-

**Table 7.** Charged groups of human  $\alpha$ -thrombin that form charge patches

Charge patch	Residue	Interaction energy (kcal/mol)	Nearest charged neighbor (distance and interaction energy) Å/residue/(kcal/mol)
<b>Positive patches</b>			
p1	Lys 36	0.3	(6.7/Lys 109/0.4)
	Arg 73	0.6	(8.2/Lys 149E/0.3)
	Arg 75	0.0	(6.6/Glu 77/-0.5)
	Arg 77A	0.1	(>10./)
	Lys 81	0.3	(9.9/Lys 110/0.1)
	Lys 109	0.4	(6.7/Lys 36/0.4)
	Lys 110	0.2	(9.4/Lys 109/0.1)
	Lys 145	-0.2	(>10./)
	Lys 149E	0.2	(6.6/Asp 21/-0.3)
p2a	Arg 93	1.4	(3.6/Arg 101/1.3)
	Arg 101	0.4	(3.6/Arg 93/1.3)
	Arg 126	0.6	(4.3/Lys 236/0.4)
	Arg 233	0.3	(4.9/Asp 178/-0.7)
	Lys 235	-0.5	(6.8/Asp 125/-0.5)
	Lys 236	0.7	(4.3/Arg 126/0.4)
	Lys 240	0.3	(9.7/Lys 236/0.1)
p2b	Arg 165	0.2	(4.5/Lys 169/0.8)
	Lys 169	0.3	(4.5/Arg 165/0.8)
p2c	Arg 97	0.0	(>10./)
	Arg 173	1.0	(8.6/Glu 217/-0.2)
	Arg 195	1.0	(4.1/Glu 97A/-0.9)
<b>Negative patches</b>			
n1a	Glu 39	-1.7	(3.8/Arg 35/-1.3)
	Asp 60E	-0.1	(>10./)
	Glu 61	-1.4	(4.1/Lys 87/-1.2)
	Asp 63	-0.8	(7.4/Lys 36/-0.4)
	Glu 97A	-1.1	(4.1/Arg 175/-0.9)
	Glu 146	-0.5	(3.9/Arg 221A/-1.4)
	Asp 189	1.1	(6.6/Asp 221/0.7)
	Glu 192	0.4	(7.7/Glu 146/0.4)
	Glu 217	-1.0	(3.7/Lys 224/-1.4)
	Asp 221	0.7	(4.5/Arg 187/-1.2)
n1b	Glu 18	0.4	(6.0/Asp 221/0.6)
	Glu 164	-0.9	(3.5/Lys 185/-1.5)
	Asp 170	-0.1	(8.4/Lys 169/-0.2)
	Asp 186A	0.3	(6.4/Glu 186B/0.4)
	Glu 186B	0.0	(5.2/Lys 185/-0.8)
	Asp 222	-0.7	(3.0/Arg 187/-1.7)
n2	Asp 1A	-2.1	(3.0/Lys 9/-1.8)
	Glu 1C	-0.3	(7.6/Arg 206/-0.5)
	Asp 49	-1.0	(4.0/Arg 50/-1.5)
	Glu 247	-0.8	(4.0/Arg 50/-1.5)
	CTR <sup>a</sup> 247	-0.1	(5.8/Glu 247/0.5)
n3	Glu 14H	0.1	(7.4/Glu 14E/0.4)
	Asp 14L	0.3	(6.8/CTR 15/0.2)
	CTR 15	0.2	(6.8/Asp 14L/0.2)
	Glu 127	-0.1	(7.3/Asp 125/0.2)

<sup>a</sup> Carboxy-terminus.

tant role as an anchor for the distal positively charged groups of substrate P1 residues.

The aromatic residues are relatively evenly distributed over the thrombin molecule. Edge-on-plane as well as

**Table 8.** Dihedral angles of the disulfide bridges in thrombin

Disulfide bridge	$\kappa 1$	$\kappa 2$	$\kappa 3$	$\kappa 4$	$\kappa 5$	Hand
Cys 1-122	68	77	97	-63	179	Right
Cys 42-58	-82	-163	-95	-78	-77	Left
Cys 168-182	-62	-61	-90	-77	90	Left
Cys 191-220	-157	46	88	-170	-55	Right

planar stacking of adjacent aromatic rings can be observed (e.g., Tyr 60A-Trp 60D and Phe 225-Trp 215). The concentration of four relatively exposed indole side chains (Trp 60D, Trp 148, Trp 96, and Trp 215) around the substrate-binding site is particularly remarkable (see below).

#### *The thrombin A-chain and its interaction with the B-chain*

The thrombin A-chain with its multiple-turn/helical conformation (see Fig. 3) is fixed to the B-chain surface opposite to the substrate-binding cleft (Fig. 2; Kinemage 3). It has an overall boomeranglike shape and nestles in such a way toward the B-chain that it merges to the common molecular surface without significant breaks (with the exception of the extending carboxy-terminal segment and the exposed elbow around Lys 10-Ser 11 and Glu 13-Arg 14D [Fig. 9]). The buried surface between the chains amounts to 1,200 Å<sup>2</sup>; this number can be compared to the intermolecular surface of 1,725 Å<sup>2</sup>, buried between  $\alpha$ -thrombin and hirudin in their complex (Rydel et al., 1991).

Besides the covalent disulfide bridge connection between Cys 1 and Cys 122 the interaction of the A- with the B-chain is mediated through charged side chains. Only three inter-main-chain hydrogen bonds exist between A and B (see Fig. 3), whereas another 20 hydrogen bonds are formed utilizing at least one side chain (Fig. 10). Six of the latter represent simultaneously (largely) buried salt bridges (see Table 6), and a charged side-chain group participates in a further 10 (such hydrogen bonds involving charged groups are particularly strong).

Stabilization within the A-chain occurs preferentially through polar and salt-bridge interactions (in particular in its central part [see Figs. 9, 10]). Fourteen of the intra-A-chain hydrogen bonds are made between main-chain groups (Fig. 3). Nine oppositely charged side-chain groups are involved in four of the six residual hydrogen bonds and form part of the extended salt-bridge clusters (Table 6); three of them contribute simultaneously to the interchain salt bridges mentioned above (Fig. 10).

The amino-terminal segment of the A-chain up to Glu 1C is only defined by relatively weak density (see Fig. 28) in the PPACK- $\alpha$ -thrombin crystals but is placed unequivocally. This segment runs along a ridge of the B-



Table 9. Internal water molecules and their hydrogen-bond partners as found in thrombin PPACK<sup>a</sup>

PPACK-human $\alpha$ -thrombin					Bovine trypsin		PPACK-human $\alpha$ -thrombin					Bovine trypsin	
Solvent number	Hydrogen-bond partners			Solvent number	Deviation of equivalent molecules (Å)	Solvent number	Hydrogen-bond partners			Solvent number	Deviation of equivalent molecules (Å)		
	Atom	Distance (Å)	B (Å <sup>2</sup> )				Atom	Distance (Å)	B (Å)				
304	Asn 143	Nδ2	(2.9)	12	—	—	325	Met 26	O	(2.7)	15	—	—
	Lys 145	O	(3.2)	—	—	—		Arg 137	NH	(3.0)	—	—	—
	Thr 147	N	(3.2)	—	—	—		Sol 454	OH	(2.7)	—	—	—
305	Trp 215	O	(3.4)	4	416	0.9	327	Pro 5	N	(3.5)	31	—	—
	Phe 227	O	(3.3)	—	—	—		Leu 6	N	(3.1)	—	—	—
	Arg 31	NH1	(3.0)	—	—	—		Gly 25	O	(2.6)	—	—	—
	Tyr 228	OH	(3.7)	—	—	—		Asp 116	O	(2.8)	—	—	—
306	Met 210	O	(3.0)	14	722	0.7	332	Arg 101	N	(2.9)	28	703	0.3
	His 230	Nδ1	(3.2)	—	—	—		Asp 102	N	(3.5)	—	—	—
	Phe 232	N	(2.9)	—	—	—		Asn 179	O	(2.9)	—	—	—
308	Ser 45	Oγ	(3.1)	13	406	0.7		Sol 368	OH	(2.7)	—	—	—
	Leu 53	O	(2.8)	—	—	—	334	Ser 171	Oγ	(3.0)	15	—	—
	Gly 196	O	(2.6)	—	—	—		Tyr 225	N	(2.9)	—	—	—
	Gln 209	Nε2	(2.9)	—	—	—		Sol 411	OH	(2.9)	—	—	—
310	Glu 217	O	(2.7)	8	415	0.3	335	Trp 148	Nε1	(2.9)	23	—	—
	Gly 219	O	(3.5)	—	—	—		Glu 192	Oε1	(3.1)	—	—	—
	Cys 220	N	(3.4)	—	—	—		Cys 220	N	(3.4)	—	—	—
	Arg 221	N	(3.0)	—	—	—		Cys 220	Sγ	(3.2)	—	—	—
311	Trp 141	N	(3.1)	19	410	0.4	337	Glu 217	N	(2.9)	22	—	—
	Gly 193	O	(2.8)	—	—	—		Tyr 225	O	(2.6)	—	—	—
	Sol 314	OH	(3.2)	—	—	—		Sol 310	OH	(3.4)	—	—	—
314	Glu 30	Nε2	(3.2)	10	701	0.1		Sol 411	OH	(3.3)	—	—	—
	Thr 139	O	(2.7)	—	—	—	342	Arg 221	O	(2.7)	10	—	—
	Asp 194	O	(3.0)	—	—	—		Lys 224	O	(2.8)	—	—	—
	Sol 311	OH	(3.2)	—	—	—		Sol 316	OH	(2.8)	—	—	—
315	Ala 183	O	(3.0)	16	704	0.3		Sol 343	OH	(2.4)	—	—	—
	Asp 189	Oδ2	(2.8)	—	—	—		Sol 414	OH	(3.3)	—	—	—
	Gly 226	N	(3.5)	—	—	—	343	Tyr 184	O	(2.9)	10	—	—
	Tyr 228	OH	(2.8)	—	—	—		Lys 224	O	(2.6)	—	—	—
316	Asp 221	Oδ1	(2.7)	11	—	—		Sol 320	OH	(3.1)	—	—	—
	Sol 342	OH	(2.5)	—	—	—		Sol 342	OH	(2.4)	—	—	—
	Sol 320	OH	(2.8)	—	—	—		Sol 345	OH	(2.6)	—	—	—
318	Glu 30	O	(3.0)	17	708	0.2	345	Gly 188	O	(2.8)	27	705	0.5
	Arg 67	O	(2.7)	—	—	—		Sol 315	OH	(3.3)	—	—	—
	Gly 69	N	(3.3)	—	—	—		Sol 320	OH	(3.1)	—	—	—
	Lys 70	N	(3.2)	—	—	—		Sol 343	OH	(2.6)	—	—	—
319	Ile 16	N	(2.7)	3	430	0.3	359	Tyr 117	O	(3.0)	39	—	—
	Gly 140	O	(3.1)	—	—	—		Sol 358	OH	(3.3)	—	—	—
	Gly 142	N	(2.9)	—	—	—		Sol 360	OH	(2.7)	—	—	—
	Gly 142	O	(2.7)	—	—	—		Sol 457	OH	(3.3)	—	—	—
320	Tyr 184	N	(2.9)	13	—	—	360	Ser 27	O	(3.4)	15	709	1.2
	Gly 188	O	(3.0)	—	—	—		Lys 70	O	(2.8)	—	—	—
	Sol 316	OH	(2.8)	—	—	—		Sol 322	OH	(3.0)	—	—	—
	Sol 343	OH	(3.1)	—	—	—		Sol 359	OH	(2.7)	—	—	—
	Sol 345	OH	(3.1)	—	—	—		Sol 457	OH	(3.0)	—	—	—
321	Arg 137	O	(3.2)	21	721	0.4	368	Leu 99	O	(2.8)	16	408	0.4
	Thr 139	Oγ	(2.8)	—	—	—		Ser 214	Oγ	(2.9)	—	—	—
	Pro 198	O	(3.0)	—	—	—		Sol 322	OH	(2.7)	—	—	—
	Sol 360	OH	(3.0)	—	—	—	386	Val 241	O	(2.6)	21	—	—
323	Lys 185	O	(2.8)	28	—	—		Ile 242	O	(2.8)	—	—	—
	Lys 186D	O	(3.0)	—	—	—		Gly 244	N	(3.2)	—	—	—
	Sol 324	OH	(2.9)	—	—	—		Phe 245	N	(3.2)	—	—	—

(continued)



Table 9. Continued

PPACK-human $\alpha$ -thrombin				Bovine trypsin		
Solvent number	Hydrogen-bond partners		B ( $\text{\AA}^2$ )	Solvent number	Deviation of equivalent molecules ( $\text{\AA}$ )	
	Atom	Distance ( $\text{\AA}$ )				
410	Arg 165	N $\epsilon$	(2.7)	25	—	—
	Arg 165	NH2	(3.2)			
	Thr 177	O	(3.0)			
	Met 180	O	(3.1)			
411	Glu 217	O $\epsilon$ 2	(3.0)	25	—	—
	Sol 334	OH	(2.9)			
	Sol 337	OH	(3.3)			
414	Tyr 184	O	(2.7)	21	—	—
	Gly 223	N	(2.7)			
	Sol 342	OH	(3.3)			
454	Ser 27	O $\gamma$	(3.3)	25	—	—
	Trp 29	N $\epsilon$ 1	(3.0)			
	Sol 321	OH	(2.9)			
	Sol 325	OH	(2.7)			
457	Ser 27	N	(2.9)	35	717	0.7
	Ser 27	O	(3.3)			
	Sol 359	OH	(3.3)			
	Sol 360	OH	(3.0)			
626	Trp 51	N $\epsilon$ 1	(3.0)	44	713	0.8
	Lys 107	N $\zeta$	(3.4)			

<sup>a</sup> Topologically equivalent solvent molecules of bovine trypsin (benzamidine-free, DEBA [Bode et al., 1976]) are added together with the r.m.s. deviation obtained after optimal superposition of both enzyme models (for deviations up to 1.2  $\text{\AA}$ ).

chain with only a few van der Waals contacts between both chains. The position of the amino-terminal Thr 1H is almost in the center of the carboxy-terminal (nonhelical) half-turn of the B-chain (see Fig. 9). It is close to the negatively charged side chains of Asp 243 and Glu 247 but not within direct hydrogen bond distance (see Table 5).

The equivalent A-chain segment seems to be similarly arranged in bovine thrombin; in the crystals of human thrombin bound to hirudin (Rydel et al., 1991) it is definitely in a quite different position and conformation, presumably due to intermolecular contacts. The A-chain amino-terminus therefore possesses a high degree of flexibility and can adapt to different crystal environments but seems inherently to prefer the location found in PPACK-thrombin. The preceding segment, cleaved off autocatalytically upon activation in human prothrombin, might extend away from the molecular surface thus facilitating enzymatic attack; in bovine thrombin the equivalent 13-residue segment, which is not cleaved off in the course of activation, is completely undefined by electron density, indicating its complete disorder.

The central, more rigid part of the A-chain (between Asp 1A and Tyr 14J) runs in a shallow curved groove of the B-chain lined by B-chain segments Ser 20–Met 26 and Lys 202–Trp 207 (Fig. 9). The amino-terminal half of this central segment, organized mainly in tightly packed turns, is stabilized considerably by several intrachain and interchain hydrogen bonds (Fig. 3) and by salt-bridge clusters (Table 6). Asp 14 and Lys 202, with their side chains pointing toward the molecular center and toward the peripheral A-chain, respectively, are at the center of the charge network (Fig. 10). The carboxylate group of Asp 14 opposes the guanidyl group of (B-chain residue) Arg 137 almost symmetrically; furthermore, it is involved in favorable hydrogen bond/salt bridge interactions with the guanidyl group of (A-chain residue) Arg 4. Glu 8 makes hydrogen-bond/electrostatic interactions with this Arg 4, but also with Lys 202, which further interacts with the carboxylate group of Glu 14C.

Most of these salt bridges are more or less completely buried (Table 6). Electrostatic interactions therefore contribute significantly to the intrachain stability as well as to the A–B interaction; almost 90% of the total electrostatic energy of the interaction between A and B is calculated to be due to these salt clusters. Six out of 7

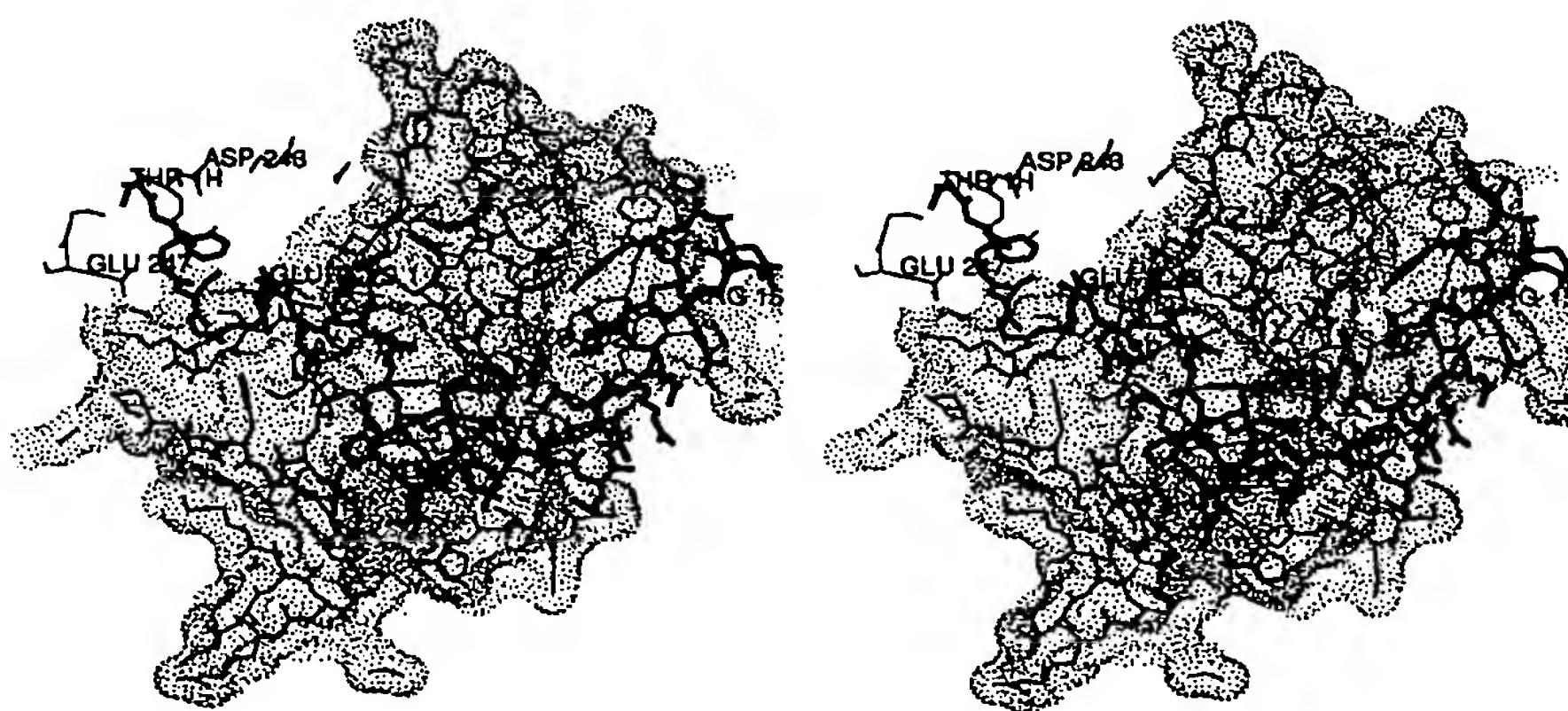


Fig. 9. Back view of the human thrombin B-chain (thin connections) and its Connolly surface, displayed together with the thrombin A-chain (thick connections). This figure is obtained from Figure 1 after an approximately 180° rotation about a vertical axis. The A-chain runs from left (close to the B-chain carboxy-terminus Glu 247) to right along a boomeranglike shallow groove of the B-chain.

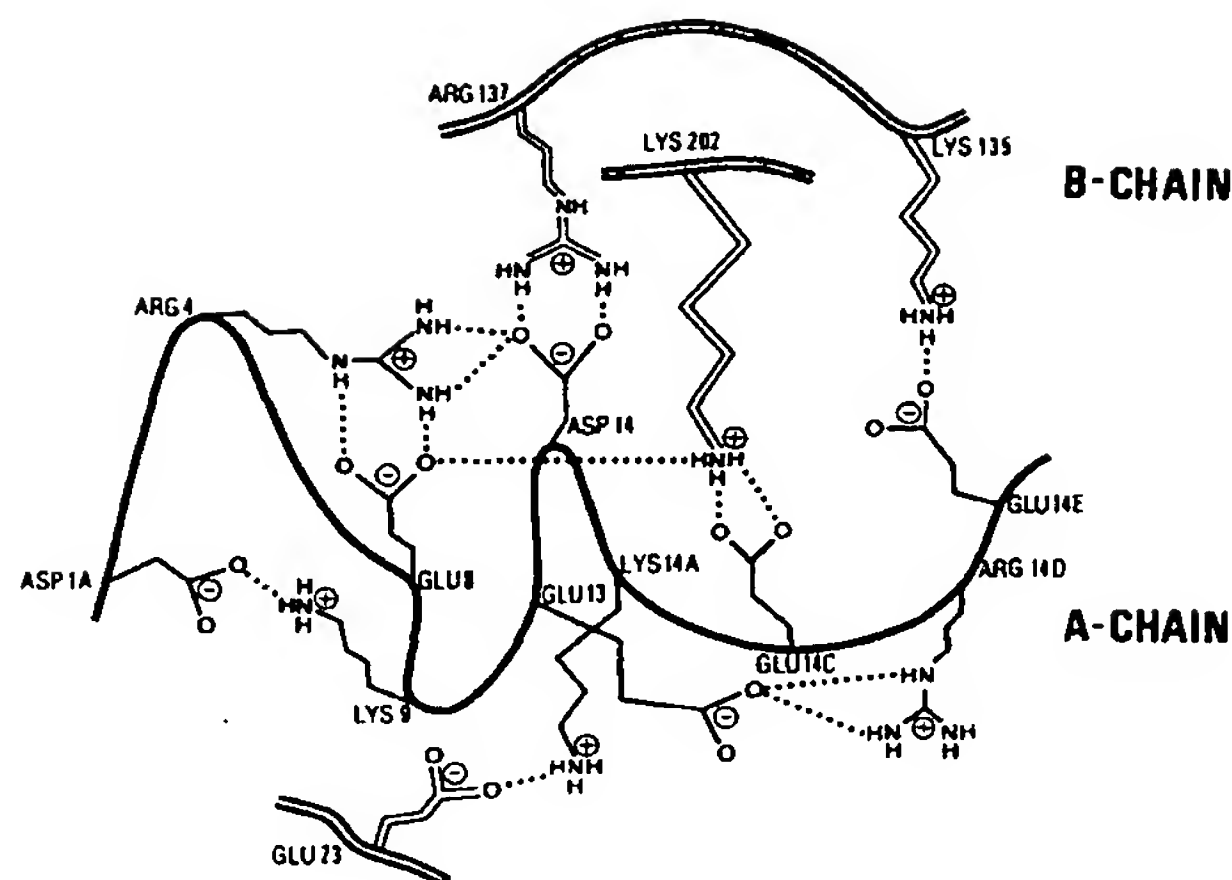


Fig. 10. Schematic drawing of the ionic interactions (as far as involved in hydrogen bonding) within the thrombin A-chain and between the thrombin A-chain (bold connections) and the B-chain (double connections). The arrangement is similar to that in Figure 9, i.e., corresponds to a back view of the thrombin molecule.

A-chain residues involved in intra- and interchain salt bridges are shared by human and bovine thrombin; only Glu 13, which makes a weak surface-located salt bridge in human thrombin, is replaced by a glutamine in the bovine species. In spite of these impressive electrostatic interactions, however, the A-chain is apparently neither of great importance for B-chain folding nor for the (thermodynamic) stability of the B-chain globule (Hageman et al., 1975).

The carboxy-terminal A-chain segment between Thr 14B and Tyr 14J is organized in an amphiphilic  $\alpha$ -helix of  $1\frac{1}{2}$  turns (Figs. 3, 9). The side chains of Leu 14G and Tyr 14J are involved in hydrophobic contacts between the A- and the B-chains. The phenolic side chain of Tyr 14J exhibits enhanced flexibility. The segment from Ile 14K

to the carboxy-terminal Arg 15, which projects away from the main molecular body, is only weakly defined by electron density. The carboxy-terminus at Arg 15 is positioned close to the guanidyl group of Arg 165 of a symmetry-related molecule. Segment Ile 14K–Arg 15 is arranged quite differently in bovine thrombin as well as in the human thrombin–hirudin complex; it therefore appears to be inherently flexible.

A-chain segment Cys 1–Leu 6 is topologically equivalent to the first six amino acid residues of the activation peptide of chymotrypsin(ogen) and connected in a similar manner via disulfide bridge Cys 1–Cys 122 to the thrombin B-chain. The more carboxy-terminal part of the thrombin A-chain is of a different conformation and considerably longer than the corresponding pro-part of chymotrypsinogen. The interaction surface of the thrombin B-chain has a relatively similar contour to that of chymotrypsinogen; in thrombin, however, most of the interacting side chains are longer and more polar (see Fig. 11).

In human  $\alpha$ -thrombin, Arg 15 is about 20 Å away from the equivalent position in chymotrypsinogen (Fig. 11). Assuming an A-chain conformation identical to  $\alpha$ -thrombin up to Tyr 14J, Arg 15 of a prothrombin molecule would not be able to attain an equivalent position, even after a complete extension of the connecting polypeptide segment. Presumably, therefore, thrombin segment Ile 16–Ser 20 occupies a different position in prothrombin.

As mentioned above, the B-chain alone appears to be thermodynamically stable and to retain clotting activity (Hageman et al., 1975). Thus, the A-chain would not seem to be a structural element required for the activity of thrombin; its integration into the B-chain might rather be necessary for correct activation cleavage and for other functions of prothrombin. As it is positioned opposite to the active site, interaction of the A-chain with larger bound substrates and inhibitors is not possible. It is, however, conceivable that amino-terminally linked pro-parts,

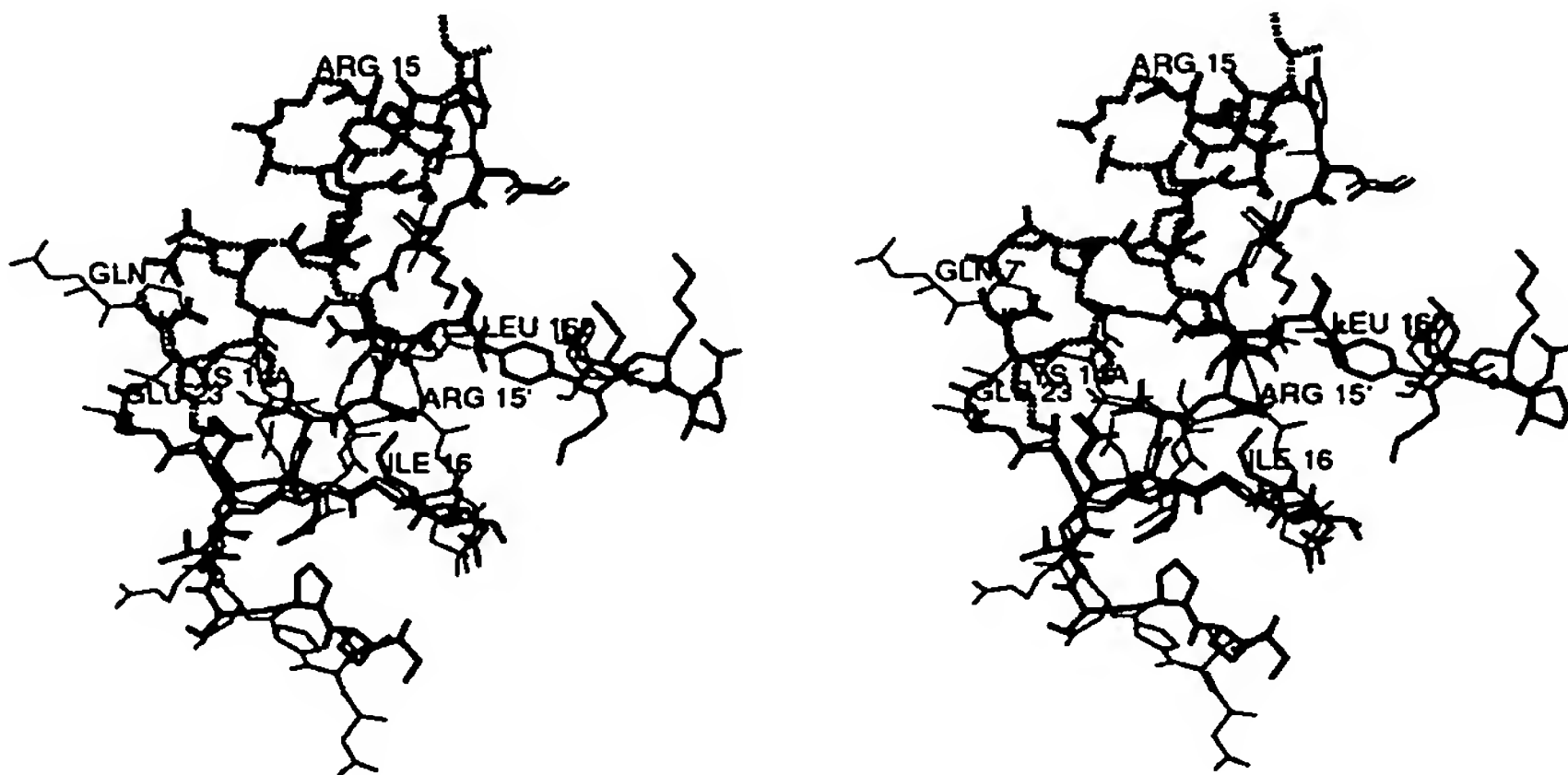


Fig. 11. Parts of the thrombin A-chain (dashed connections) and B-chain (thick connections) optimally superimposed with equivalent parts of bovine chymotrypsinogen (thin connections). Polypeptide segment 10–20 of chymotrypsinogen deviates considerably from the equivalent parts of both thrombin chains; the carboxy-terminal residue Arg 15 of the thrombin A-chain is quite distant from the equivalent Arg 15' of the (nonactivated!) chymotrypsinogen. This back view is similar, but not identical, to that used in Figure 9.

such as fragment 1.2 of meizothrombin, could fold back and project into the substrate-binding region, thus impairing activity toward macromolecular substrates, as recently shown for fibrinogen or Factor V (Doyle & Mann, 1990).

### The thrombin B-chain

The tertiary structure of the thrombin B-chain is homologous to the reactive domains of the other trypsin-like serine proteinases of known spatial structure (see Figs. 2, 4, 5 and Bode et al., 1989b). A remarkable feature of the thrombin B-chain globule is, however, a series of more elongated and exposed loops, in particular around the active-site cleft. The unique specificity of thrombin for macromolecular ligands involves interactions at the surface of the molecule. The surface segments of the thrombin B-chain exhibit new features compared with chymotrypsinogen (as can be inferred from Table 3, Figs. 4, 5 and as has been presented in more detail by Bode et al., 1989b). Among them are (1) segment Arg 173-Ile 174; (2) loop segment 95-100; (3) the large insertion loop Leu 59-Asn 62 (the 60-insertion loop or Trp 60D loop), and (4) loop segment 34-42, which together form the "north rim" of the active-site canyon (Bode et al., 1989b); (5) the large insertion loop 145-150, which together with the trypsinlike segment Gly 216-Cys 220 forms the corresponding "south rim"; (6) loop segment Tyr 184A-Gly 188, and (7) the trypsinlike loop segment Cys 220-Tyr 225, which together make up the thrombin structure behind the specificity pocket; and (8) loop segment Ser 203-Arg 206, which is particularly involved in the interaction with the A-chain (see above).

Most of these surface loops project out of the molec-

ular surface. However, with the exception of loop segment 145-150 (see above and below), all of these loops are remarkably rigid. Proline and tryptophan residues seem to contribute to their rigidity.

The positions of some polypeptide segments of the  $\alpha$ -thrombin B-chain particularly susceptible to proteolytic cleavage are emphasized in Figure 12 and Kinemage 2. Some of these cleavage sites (in particular those generated autocatalytically or through trypsin action) may only exist under laboratory conditions and might play no important role in vivo; they are, nevertheless, extremely valuable for investigating the importance of distinct thrombin sites for various thrombin functions. Most of these sites are located at exposed surface loops. One of them (Lys 149-Gly 150, see below) is not readily accessible to an attacking proteolytic enzyme; another one (Arg 67-Ile 68) is even completely buried in the interior of the thrombin molecule and only accessible after prior cleavage in the 70-80 loop (see below). None of the preferred proteolytic cleavage sites shown in Figure 12 exhibit the canonical conformation found in substratelike binding small protein inhibitors (Huber & Bode, 1978; Laskowski & Kato, 1980) (see Table 10). Most of these sites are, however, flexible enough to allow conformational adaptation to the binding sites of the proteolytic enzymes. The rigid, noncomplementary activation cleavage site of chymotrypsinogen (Wang et al., 1985), however, illustrates nicely that the susceptibility of peptide bonds is governed by other (hitherto unknown) properties.

### Thrombin sites of particular importance

#### Ile 16-Asp 194 and its role in activation

Similar to the other activatable trypsinlike serine proteinases, the amino-terminal residue Ile 16 (and part of

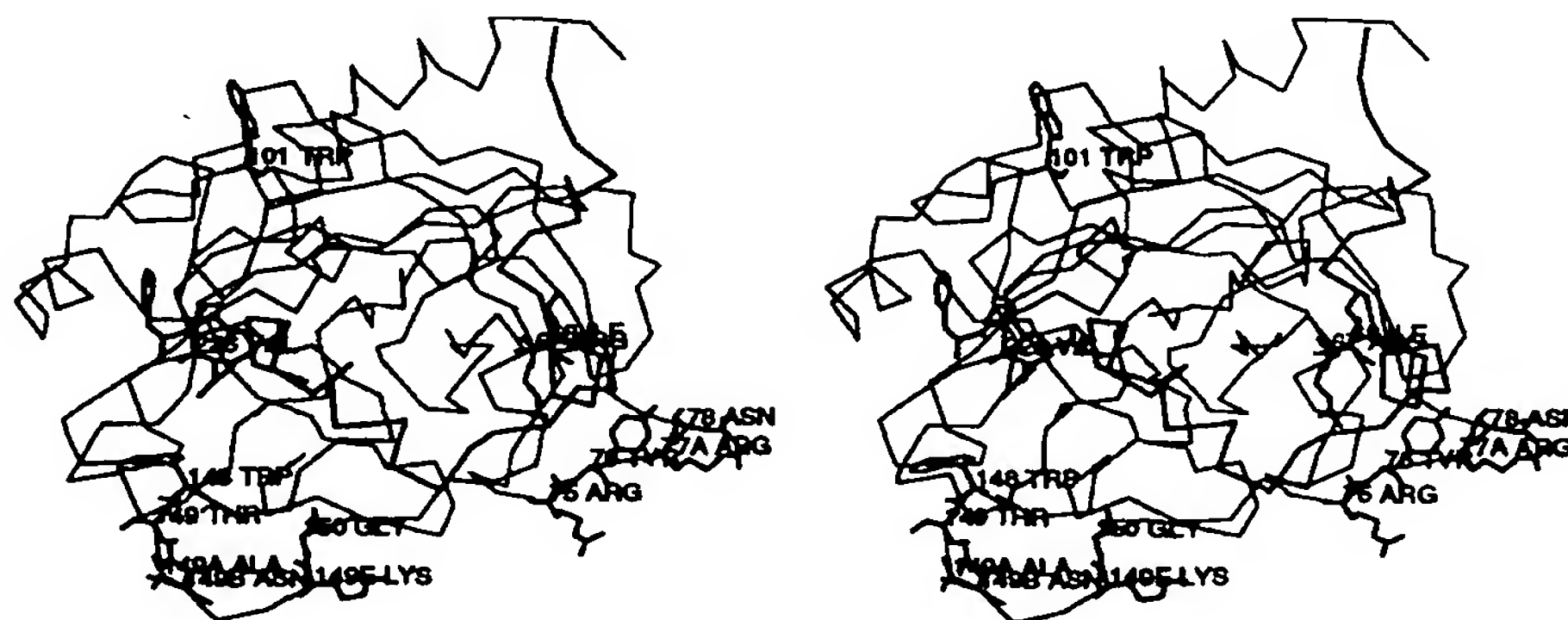


Fig. 12.  $\alpha$ -carbon chain of the A-chain (thick connections) and the B-chain (thin connections) of human  $\alpha$ -thrombin displayed together with the full amino acid residues placed at well-characterized proteolytic cleavage or replacement sites in naturally occurring thrombin variants. The bound PPACK is overlaid. Cleavages at Arg 77A-Asn 78 and Arg 67-Ile 68 give rise to  $\beta$ -thrombin; cleavages at Trp 148-Thr 149, Ala 149A-Thr 149B, and Lys 149E-Gly 150 yield  $\zeta$ -,  $\epsilon$ -, and  $\gamma$ -thrombin, respectively. In the genetic thrombin variants Quick I, Quick II, and Tokushima, Arg 67, Gly 226, and Arg 101 are replaced by cysteine, valine, and tryptophan, respectively. Standard view as in Figure 1.



**Table 10.** Main-chain conformational angles of some  $\alpha$ -thrombin cleavage sites compared with the proteinase-binding loops of some protein inhibitors around the scissile peptide bond

Scissile peptide bond	P2	P1	P1'	P2'
$\beta_1$ : Arg 77A-Asn 78	-98/95	-48/-55	-115/29	-114/-52
$\beta'$ : Arg 75-Tyr 76	-124/-30	-85/151	-67/118	-98/95
$\beta$ : Arg 67/Ile 68	-130/124	-111/118	-115/143	79/8
$\gamma_1$ : Lys 149E-Gly 150	110/164	-75/149	-92/-22	-127/137
$\gamma$ : Trp 148-Thr 149	-101/-6	-73/-8	-64/128	-163/154
$\epsilon$ : Ala 149A-Asn 149B	-137/150	-64/128	-163/154	-100/153
BPTI: Arg 151-Ala 161	-82/160	-120/34	-88/179	-132/84
Eglin c/subtilisin: Leu 45-Asp 46	-61/143	-115/43	-96/169	-117/110
OMTKY3: Leu 181-Glu 191	-76/155	-99/29	-82/156	-104/106
PPACK	-74/139	-87/		

Val 17) of the thrombin B-chain is buried within the protein moiety (see Kinemage 1). The internal salt bridge made between the  $\alpha$ -ammonium group of Ile 16 and the side-chain carboxylate group of Asp 194 is of considerable strength (3.0 kcal/mol [see Table 6]). According to our electrostatic calculations, the pK of this ammonium group is shifted to about 10.5, due mainly to this strong interaction. The deprotonation of this ammonium group and the concomitant disruption of the salt bridge is generally believed to result in a transformation of the active enzyme structure toward a zymogen-like structure (for a more detailed discussion of the equivalent case in trypsin, see Bode [1979]). This transformation is probably accompanied by a decrease in activity with increasing pH values in the alkaline range. Indeed, the thrombin hydrolytic activity toward typical synthetic *p*-nitroanilide substrate seems to be governed by an ionizable group of a pK value of almost 10 (see Lottenberg et al., 1983; De Cristofaro & DiCera, 1990).

In prothrombin the intact Arg 15-Ile 16 segment must be situated on the molecular surface (see Fig. 11). It would be reasonable to assume that segment 191-194 and the adjacent residues in prothrombin might have a similar conformation as the equivalent residues in chymotrypsinogen (Freer et al., 1970; Wang et al., 1985); i.e., that Asp 194 might project into the active-site cleft and loop segment 191-194 folded inward with residue 192 occupying the specificity pocket. However, the thrombin residue topologically equivalent to His 40 of trypsinogen and chymotrypsinogen is Leu 40; thus, Asp 194 of prothrombin could not, as in the chymotrypsinogen structure, accept a (probably stabilizing) hydrogen bond from the side chain of residue 40. In addition, according to model-building experiments the polar side chain of Glu 192 of prothrombin would probably not find a favorable environment in the protein interior of prothrombin (in contrast to the equivalent methionine side chain of chymotrypsinogen). Therefore, both substitutions (at positions 40 and 192) seem to favor destabilization of the thrombin zymogen state relative to the active state. As mentioned earlier

(Bode, 1979), such an equilibrium shift toward the active-state-like conformation could facilitate formation of proteolytically active prothrombin-protein complexes such as the prothrombin-staphylocoagulase complex (see Kawabata et al., 1985).

#### *The insertion loop Leu 59-Asn 62 (60 insertion loop)*

The polypeptide segment Cys 58-Leu 64 of thrombin is nine residues longer than the corresponding segment of chymotrypsin (see Table 3; Fig. 4) and folded in a characteristic hairpin loop, which projects (in contrast to the digestive enzymes) considerably out of the molecular surface (Figs. 4, 5, 13A). This large insertion loop plays an important role in the selection of susceptible (macromolecular) substrates and inhibitors, i.e., for recognition and specificity of thrombin (Bode et al., 1989b). Segment Leu 59-Phe 60H is folded in an approximate  $3_{10}$ -helical turn, which is terminated by Glu 60E in a left-handed helical conformation (Table 2). The phenolic side chain of Tyr 60A points toward the S2 subsite, covering it, and packing tightly against Pro 21 of the bound PPACK molecule (see below). The indole moiety of Trp 60D is almost fully exposed to bulk solvent molecules. This part of the exposed loop seems to be particularly stabilized by the two intervening proline residues, which, in addition, pack against Trp 96 of an adjacent loop (see Fig. 13A and Kinemage 4). In the PPACK-thrombin crystals the whole loop (including all side chains) is well defined by electron density (Fig. 13B) and also makes intermolecular contacts with residues of the B-chain carboxy-terminus of a symmetry-related thrombin molecule. In spite of very different crystal contacts in the thrombin-hirudin (Rydel et al., 1991) and in the bovine thrombin structure, this loop is of similar structure, i.e., its internal conformation seems to be stable.

The well-defined (carbohydrate carrying) asparagine side chain of the subsequent Asn 60G extends from the loop surface away from the active-site cleft (Fig. 13A). In the PPACK-thrombin crystals there is free intermo-



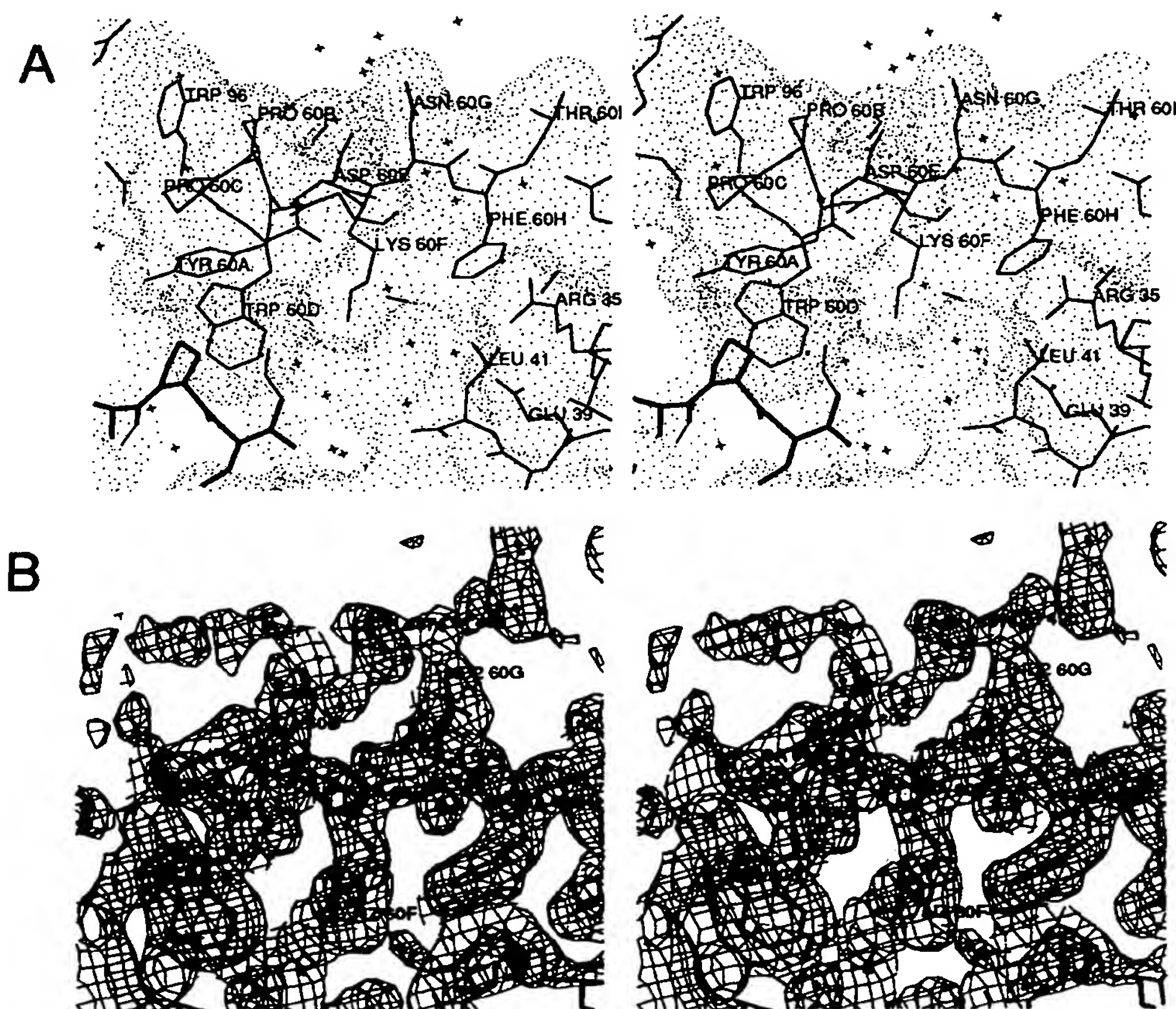


Fig. 13. A: The 60 insertion loop and the glycosylation site of human  $\alpha$ -thrombin (thin connections) displayed together with part of PPACK (thick connections) and the thrombin Connolly surface. Loop segment Tyr 60A–Asp 60E projects particularly out of the molecule; it probably forms part of the chemotactic domain of thrombin. Asn 60G is pointing away from the active-site cleft. Standard view as in Figure 1. B: Section of the final electron density around the 60 insertion loop superimposed with the equivalent parts of human  $\alpha$ -thrombin and some localized solvent molecules (crosses). The electron density representing the Asn 60G side chain (upper right corner) extends toward the top; it is, however, not appropriate to accommodate a distinct Asn 60G *N*-linked sugar chain; instead, this density has been interpreted as two fixed solvent molecules. Contour surface at 0.8 $\sigma$  calculated for the whole map. Standard view as in Figure 1.

lecular space but no continuous electron density that would indicate the carbohydrate linked to it (Fig. 13B); the isolated density blobs visible beyond the carboxamide group have therefore been interpreted as solvent molecules. The bound, presumably disordered, oligosaccharide (consisting of 12 sugar units in a branched chain [Nilsson et al., 1983]) must extend away from the active-site cleft (Fig. 13A). Thus, it should not interfere with macromolecular substrates that bind mainly along the cleft; this is in agreement with the lack of any effect on enzymatic (clotting and esterase) and cellular activities of thrombin on its removal or destruction (Horne & Gralnick, 1983). Concanavalin A, in binding to the  $\alpha$ -D-mannose of this sugar chain, might on the other hand come into conflict with an approaching fibrinogen molecule, in agreement with experimental results (Skaug & Christen-

son, 1971; Karpatkin & Karpatkin, 1974; Hageman et al., 1975).

The compact and relatively rigid segment Tyr 60A–Trp 60D is part of a tetradecapeptide with the sequence Tyr 60A–Leu 65 of known extracellular matrix-binding and chemotactic/growth factor activities (Bar-Shavit et al., 1984, 1986, 1989). The intact catalytic residues are not a prerequisite for chemotactic activity toward monocytes and neutrophils (Bizios et al., 1986). It is conceivable that such activities could be attributed to this exposed Tyr 60A–Trp 60D loop of  $\alpha$ -thrombin for its rather unique ridgelike architecture. In thrombin–hirudin this projecting thrombin loop is smoothed out by the hirudin Pro 46HI–Glu 49HI segment packing alongside of it (Rydel et al., 1991); thus, partial masking upon hirudin binding could account for the observed abolition of chemotactic

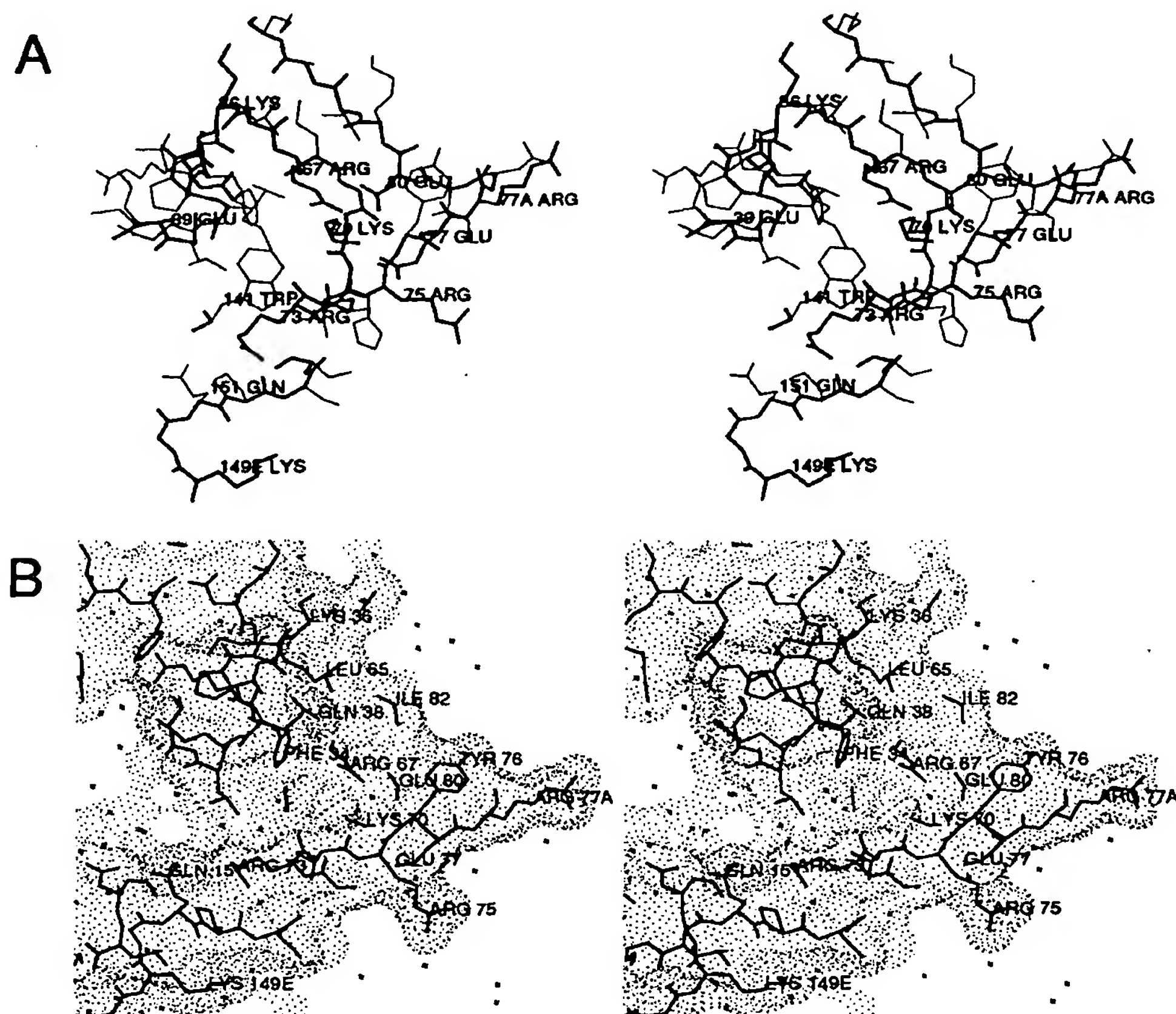
activity (Bizios et al., 1986; Prescott et al., 1990). Extracellular matrix-bound thrombin failed to form a complex with antithrombin III, but still exhibits an open proteolytic site (Bar-Shavit et al., 1989). This is in agreement with recent model-building studies, which show that the core of antithrombin III should pack against this projecting loop thus burying it and removing it from contacts with receptors (unpubl.).

According to a search in 7,068 sequences, this YPPW sequence seems (with the exception of a presumed polymerase subunit from a vesicular stomatitis virus) to be confined to thrombin. Preliminary molecular dynamics simulations with an isolated and *N*-methylated Tyr-Pro-Pro-Trp-Asn-amide peptide indicate that the distinct conformation observed in thrombin might also be a low free-energy conformation in the isolated peptide; recent NMR experiments show that the equivalent hexapeptide does, however, not possess a single preferred conforma-

tion but interchanges between at least two conformational states (unpubl.).

*Surface structure around segment 67–80  
(the fibrin(ogen) secondary binding exosite)*

The loop Lys 70–Glu 80 (Fig. 14A; Kinemage 5) of thrombin is topologically similar to the so-called calcium-binding loop of bovine trypsin (Bode & Schwager, 1975a,b) (Fig. 15A) and of the other digestive serine proteinases (see Meyer et al., 1988). The central residue of this loop is Lys 70. Its distal ammonium group occupies a site equivalent to the calcium ion in the pancreatic enzymes (Fig. 15A). In the thrombin structure, however, of the six oxygen ligands of calcium found in bovine trypsin only three are connected to Lys 70N $\epsilon$  through short hydrogen bonds, in an almost tetrahedral geometry (Table 11). It is noteworthy, that the distal ammonium group of Lys 70, which is well below the molecular surface in



**Fig. 14.** The fibrinogen secondary binding exosite of human  $\alpha$ -thrombin. The view is similar to the standard view of Figure 1. **A:** Stick drawing of loop 67–80 and its environment. Lys 70 is located at the center of an extended buried salt bridge network made with Glu 77, Glu 80, and Arg 67. Cleavage of peptide bond Arg 77A–Asn 78 leads to formation of noncoagulant  $\beta_1$ -thrombin, in which this loop is presumably unfolded. **B:** Surface residues of the exosite displayed together with the Connolly surface.

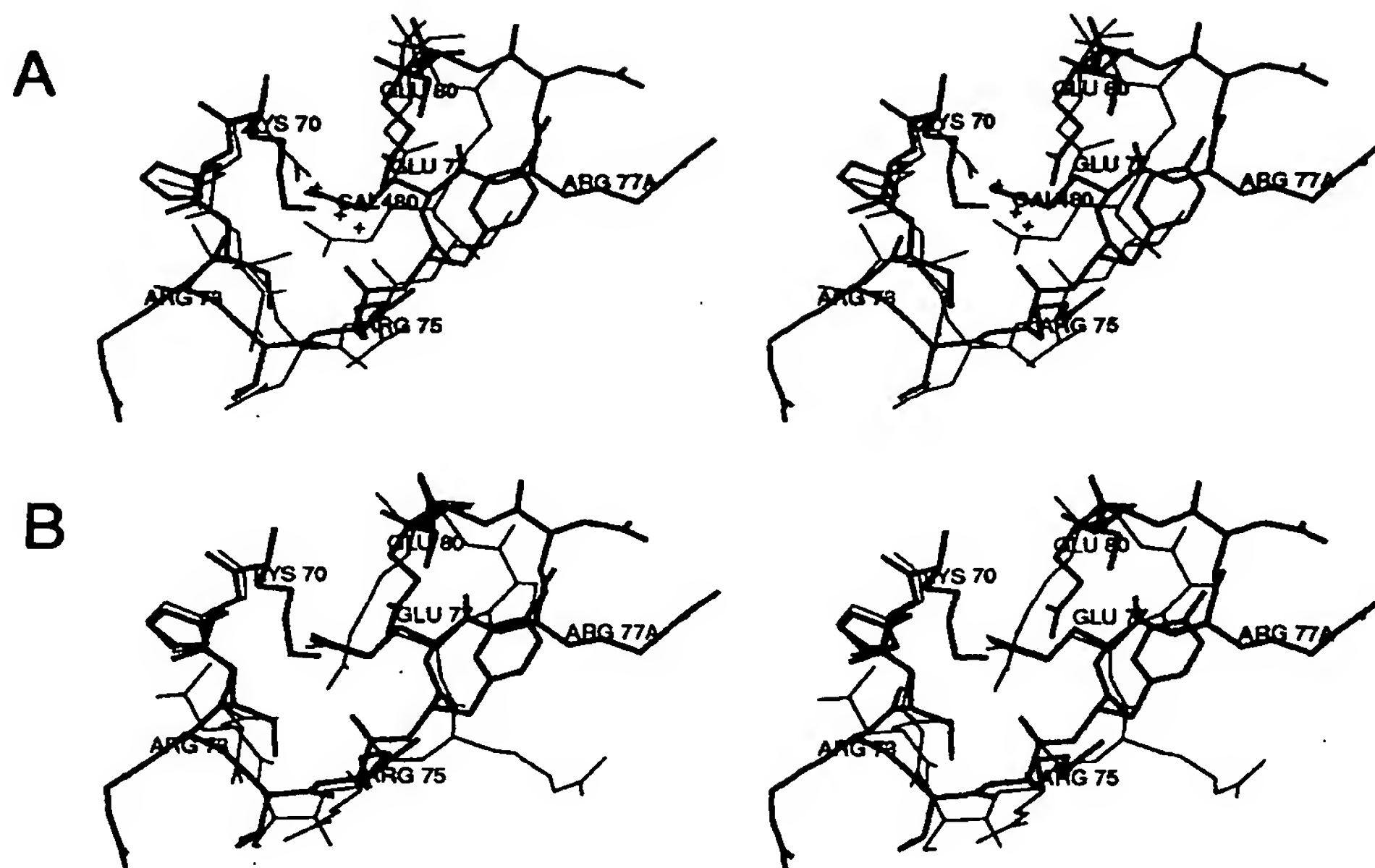


Fig. 15. A: Loop Lys 70-Glu 80 of human  $\alpha$ -thrombin (thick connections) superimposed by the calcium-binding loop, the bound calcium ion, and two adjacent solvent molecules (crosses) of bovine trypsin (thin connections [Bode & Schwager, 1975a,b]). N $\zeta$  of Lys 70 of thrombin occupies an equivalent position as the calcium ion in trypsin. The view is similar to the standard view of Figure 1. B: Loop Lys 70-Glu 80 of human  $\alpha$ -thrombin (thick connections) superimposed by the equivalent loop of human leukocyte elastase (thin connections [Bode et al., 1986]). The guanidinium group of Arg 80 of the elastase occupies a similar position as Lys 70N $\zeta$  of thrombin.

the static  $\alpha$ -thrombin structure (Fig. 14B), is (in the absence of shielding proteins such as hirudin) highly susceptible to chemical modification (Chang, 1989; Church et al., 1989). Clearly, ease and amount of modification are not necessarily reliable indicators for surface accessibility but might correlate with other molecular properties.

In turn, residue Glu 80 is involved in another salt bridge/hydrogen bond with the guanidyl group of Arg 67. Thus, the charged groups of four residues (Arg 67, Lys 70, Glu 77, and Glu 80) are cross-connected with one another to form a salt-bridge cluster essentially buried in the molecule (Fig. 14A). Its contribution to the rigidity and integrity of this loop appears considerable (see Table 6). As shown previously for trypsin (Bode, 1979), the integrity of this loop structure confers (thermal) stability to the whole molecule. Thrombin (like porcine kallikrein [Bode et al., 1983] and human leukocyte elastase [Bode et al., 1986; Fig. 15B; Table 11]) seems to carry its endogenous cationic group (70N $\zeta$ ) to order the 70-80 loop and to maintain a distinct surface contour (Fig. 14A,B); thus, these three proteinases share similar 70-80 loops with respect to conformation and a central basic residue.

As already discussed (Table 7) and shown further in Figure 14B, the slightly notched surface arching over this loop has only positively charged amino acid side chains; in particular Arg 73, Arg 75, and Arg 77A are sur-

rounded exclusively by other positively charged residues such as Arg 35, Lys 149E, Lys 81, Lys 110, Lys 109, and Lys 36, which are themselves quite distant to more peripheral negatively charged residues on the surface. Their positive charge is partially compensated by the negative charges of Glu 77 and Glu 80 involved in the ionic cluster beneath the surface (see above and Fig. 14); nevertheless, they give rise to a strong positive field around this surface (Fig. 8; Table 7).

The crystal structure analysis of the thrombin-hirudin complex (Rydel et al., 1990) revealed that this positively charged thrombin surface patch interacts with the extended hirudin tail segment (Fig. 16). Beside these distinct electrostatic interactions of hirudin, the complementary fit of hydrophobic groups seems to be of particular importance (see also Stone et al., 1987; Mao et al., 1988; Maraganore et al., 1990). There is much evidence that this thrombin exosite is also involved in thrombin binding to fibrinogen (Hageman & Scheraga, 1974; Van Nispen et al., 1977; Hogg & Blombäck, 1978; White et al., 1981; Hofsteenge & Stone, 1987; Fenton et al., 1988; Lundblad et al., 1988; Church et al., 1989; Hogg & Jackson, 1989), to fibrin and its E-domain (Liu et al., 1979; Berliner et al., 1985; Kaminski & McDonagh, 1987; Kaczmarek & McDonagh, 1988; Vali & Scheraga, 1988), to thrombomodulin (Hofsteenge et al., 1986; Hofsteenge &



**Table 11.** Distances between the central cationic group and the ligands in the crystal structures of human  $\alpha$ -thrombin (HTHR), bovine trypsin (BTRY; entry DEBA in the Protein Data Bank [Bode & Schwager, 1975a,b; Bernstein et al., 1977], human leukocyte elastase (HLE) [Bode et al., 1986], and porcine pancreatic kallikrein (KK) [Bode et al., 1983]

Central cationic group	Liganding group	Distance (Å)
<b>HTHR</b>		
Lys 70N $\zeta$	..... Arg 75 O	2.6
	..... Glu 77 OE2	2.6
	..... Glu 80 OE1	2.7
<b>BTRY</b>		
Ca <sup>2+</sup>	..... Glu 70 OE1	2.3
	..... Asn 72 O	2.3
	..... Val 75 O	2.3
	..... Glu 80 OE(1)	2.3
	..... Sol.711 OH	2.3
	..... Sol.714 OH (with connection to Glu 77)	2.3
<b>HLE</b>		
Arg 80NE NH2	..... Glu 77 OE	3.1
	..... Asn 72 O	3.1
	..... Arg 75 O	2.9
<b>KK</b>		
Arg 70 guanidyl group (weakly defined)	..... 72 O	—
	..... 75 O	—
	..... Glu 77 OE	—

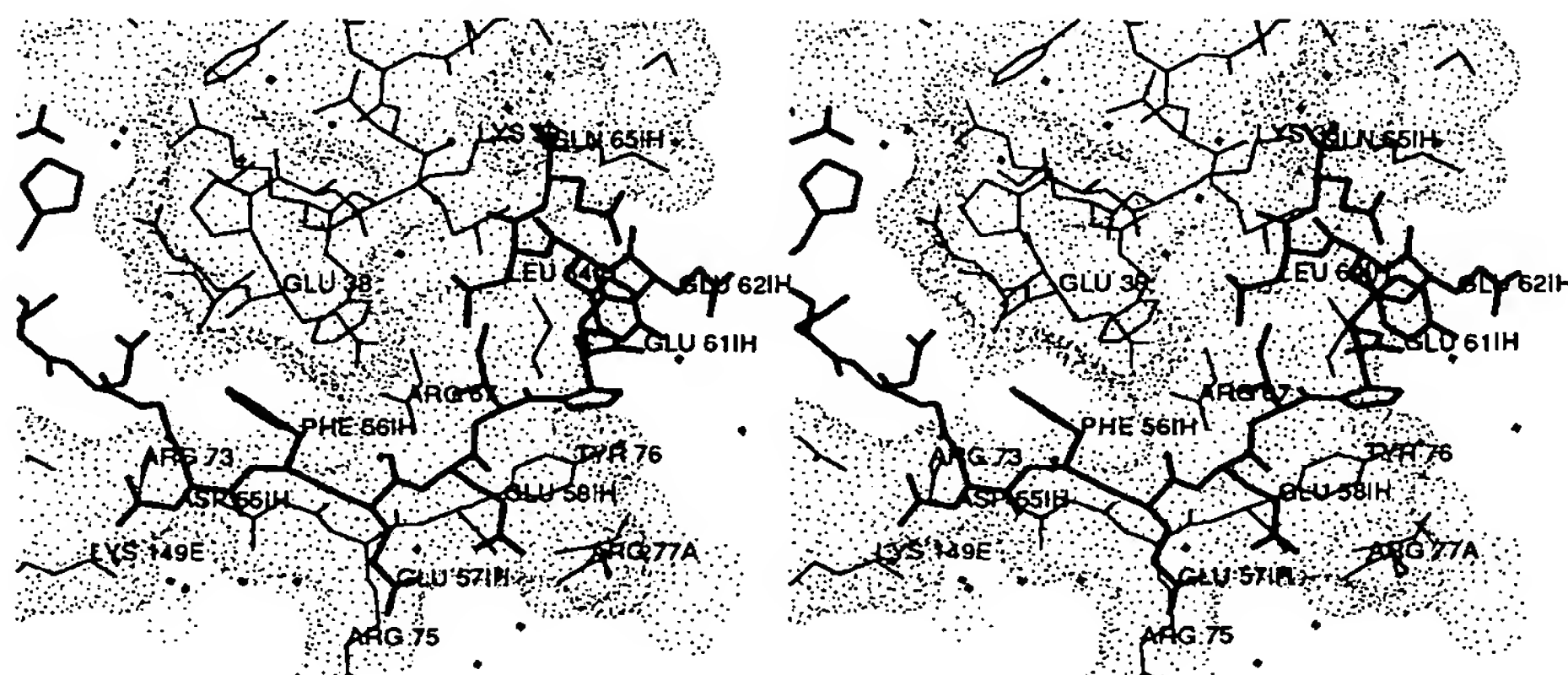
Stone; 1987; Preissner et al., 1987; Suzuki et al., 1990), and to anionic matrices (see, e.g., Fenton et al., 1988, 1989). Very recently, it has been shown that this exosite is also involved in interaction with a negatively charged

peptide segment of the platelet-derived thrombin receptor (Liu et al., 1991).

It has been shown that the above proteins bind competitively (Hofsteenge et al., 1986), i.e., that they probably occupy similar patches of this positively charged exosite of thrombin. Segment 30–44 of the fibrinogen A $\alpha$ -chain, which comprises several negatively charged residues, seems to be particularly important for binding and acceleration of the activation cleavage (Blombäck et al., 1977; Van Nispen et al., 1977; Hofsteenge et al., 1986). This fibrinogen segment exhibits only a very weak sequence homology with the hirudin tail segment, however. The thrombin–hirudin complex (Fig. 16) can, nevertheless, serve as a model for the interaction of thrombin with fibrinogen.

The positive field centered at this exosite (Fig. 8) spreads far into the extramolecular space; consequently, approaching macromolecules (substrates as well as inhibitors) could be preoriented by the combined influence of this basic exosite and the negatively charged patch around the specificity pocket (see below). Anionic groups of associated proteins need not necessarily interact through direct salt bridges with the cationic thrombin groups; they nevertheless experience the effect of this positive field at some distance from the thrombin surface. Due to its interaction with various anionic compounds and protein segments, this exosite is usually referred to as the anion-binding exosite (Fenton, 1986). This could, however, lead to confusion with the other principal basic site described below. More appropriate designations would therefore be the terms anion-binding exosite I or fibrin(ogen) secondary binding exosite (Fenton, 1981), reflecting the importance of this exosite in fibrinogen clotting and fibrin binding.

The exposed peptide bond Arg 77A–Asn 78, in the 70–



**Fig. 16.** The fibrinogen secondary binding exosite of thrombin (thin connections) and the carboxy-terminal hirudin segment (thick connections) of thrombin–hirudin (Rydel et al., 1991), displayed together with the Connolly surface of thrombin. The view is similar to the standard view of Figure 1.



80 loop (Fig. 14A), is particularly susceptible to tryptic and to autocatalytic attack (Fenton et al., 1977; Boissel et al., 1984; Chang, 1986; Bezeaud & Guillin, 1988) giving rise to  $\beta_T$ -thrombin (Braun et al., 1988). This polypeptide segment is relatively mobile (in particular from Thr 74 to Ile 79; see Fig. 28). Its main-chain conformation is quite different from the canonical conformation of typical serine proteinase protein inhibitor binding loops (Huber & Bode, 1978; Laskowski & Kato, 1980; Read & James, 1986; Bode & Huber, 1991) (see Table 10). Human  $\beta$ -thrombin is formed upon its cleavage and following excision of peptide Ile 68–Arg 77A.

Experimental evidence (Noé et al., 1988) suggests that the whole loop structure unfolds upon cleavage, driven presumably by favorable interactions of charged residues (previously involved in the buried salt cluster) with bulk water. Such a disruption of the exosite surface would explain the tremendous reduction in clotting activity and loss of affinity for fibrin, thrombomodulin, and hirudin, observed for  $\beta_T$ -thrombin (Braun et al., 1988; Hofsteenge et al., 1988) and  $\beta$ -thrombin (Fenton et al., 1977; Hofsteenge & Stone, 1987; Lewis et al., 1987; Stone et al., 1987; Bezeaud & Guillin, 1988). Indirect evidence for such an unfolding is that peptide bond Arg 67–Ile 68 (Fig. 14A) is rather susceptible to autocatalytic cleavage (Boissel et al., 1984). Being deeply buried, a cleavage of the surface Arg 77A–Asn 78 bond (see Fig. 14A,B) would be necessary prior to a proteolytic attack at this site to give  $\beta$ -thrombin. As witnessed by the almost unchanged kinetic parameters of  $\beta$ -thrombin toward chromogenic substrates this loop and surface disintegration does not seem to affect the geometry at the active site.

In the naturally occurring human genetic variant Quick I, Arg 67 (see Fig. 12) is replaced by a cysteine residue (Henriksen & Mann, 1988). The surface location of the Arg 67 guanidyl group (Fig. 14B) and its integration in an extended internal salt-bridge cluster in normal human  $\alpha$ -thrombin suggests that the exosite of this mutant might

be disrupted in a similar manner as in  $\beta$ -thrombin. This agrees with the experimental observations of tremendous loss in clotting activity but with retention of hydrolytic activity toward small synthetic substrates; the considerable reduction of platelet stimulation capability found for this mutant is a hint that this exosite is involved in platelet aggregation.

#### *Insertion loop Leu 144–Gly 150 (149 insertion loop)*

The polypeptide segment between residues 144 and 150 is five amino acid residues longer in thrombin than in chymotrypsin (see Table 3). The exposed loop (Fig. 17) lacks any inter-main-chain hydrogen bonds, but seems to be stabilized through hydrogen bonds made with its own side chains (Lys 145 and Asn 149B); some hydrophobic contacts with a symmetry-related molecule seem to further stabilize this loop in the PPACK–thrombin crystals. This loop segment attains a very different conformation and overall shape in thrombin–hirudin (Fig. 17) and is of variable conformation in the various bovine thrombin structures. In the PPACK–thrombin conformation, the Trp 148 side chain would collide with parts of the bound hirudin molecule (Rydel et al., 1991). The loop conformation observed in PPACK–thrombin may be enforced by crystal or inhibitor contacts. There is considerable evidence (e.g., from the bovine thrombin crystals) that this loop exhibits a significant degree of overall flexibility.

Three proteolyzed human thrombin variants have been found and characterized cleaved in this exposed, flexible loop segment (displayed in Fig. 12), namely (1)  $\xi$ -thrombin, resulting from cathepsin G cleavage at Trp 148–Thr 149 (Brezniak et al., 1990); (2)  $\epsilon$ -thrombin, generated by elastase action on Ala 149A–Asn 149B (Kawabata et al., 1985; Brower et al., 1987); and (3)  $\gamma$ -thrombin, resulting from tryptic or autocatalytic cleavage at Lys 149E–Gly 150 (Bing et al., 1977; Boissel et al., 1984).

As this loop is inherently flexible, the first two cleav-

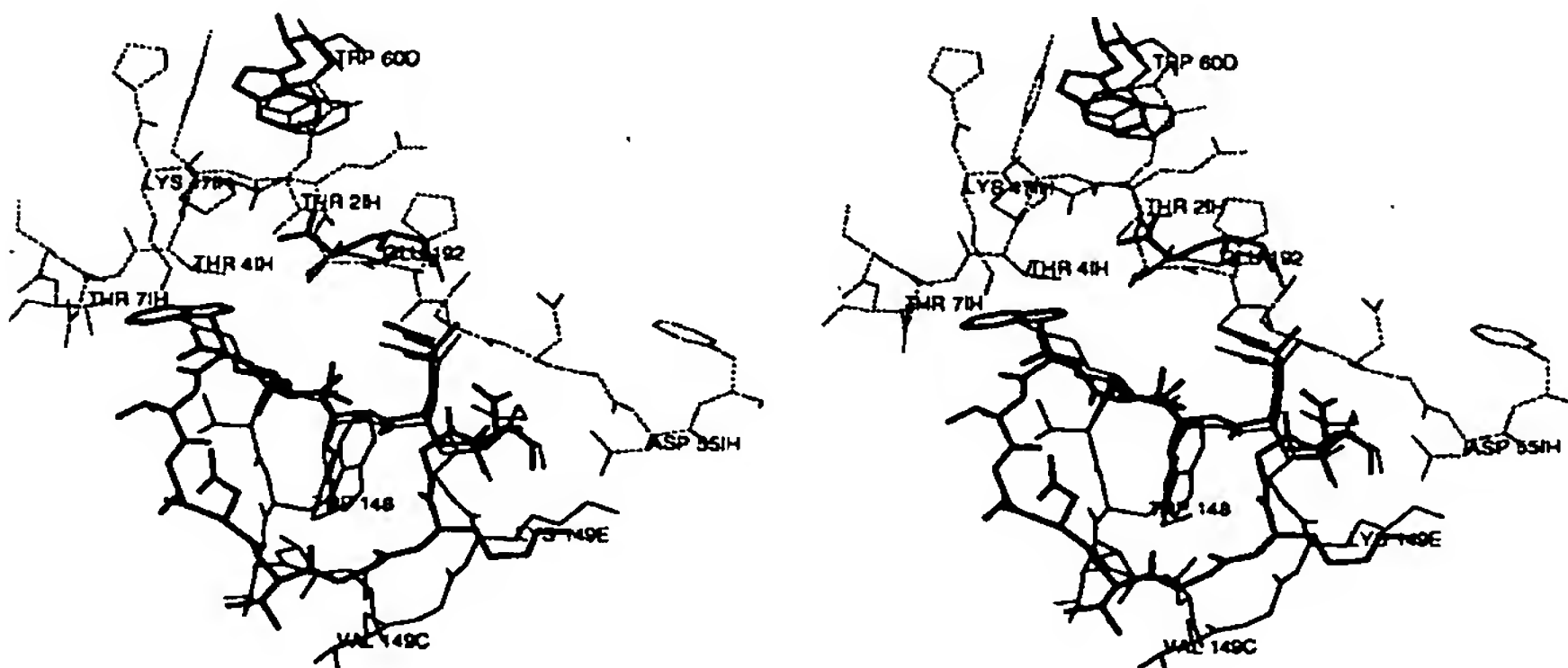


Fig. 17. Segment Asn 143–Gln 151 and Glu 192 of PPACK–thrombin (thick connections) superimposed with the same residues of the thrombin component (thin connections) and adjacent segments of hirudin (dashed connections) of the thrombin–hirudin complex (Rydel et al., 1991). Thrombin residue Trp 148 is very differently located in both thrombin molecules, and loop Glu 146–Gly 150 is differently arranged.

ages should not greatly affect the active-site geometry and thus the catalytic activity of thrombin, as demonstrated experimentally (Stone et al., 1987; Hofsteenge et al., 1988). Nevertheless, an intact loop might confer some extra stability to the whole molecule.

The last of these scissile peptide bonds, at Lys 149E–Gly 150 (Fig. 17), is somewhat obstructed in PPACK- $\alpha$ -thrombin and not readily accessible to attacking protease molecules. A primary autolytic cleavage seems possible at this site, however (Chang, 1986), probably due to the inherent mobility of this loop. This scissile peptide bond is adjacent to segment Gln 151–Pro 152, which forms (together with Glu 39 of the opposing loop) the entrance to the fibrin(ogen) secondary binding exosite (see above and Fig. 14B).  $\epsilon$ -Thrombin, cleaved four bonds before Lys 149E–Gly 150 (see Fig. 12), shows only a minor decrease in clotting activity compared with  $\alpha$ -thrombin (Hofsteenge et al., 1988). The nearly complete lack of clotting activity of  $\gamma_T$ - and  $\gamma$ -thrombin (Bing et al., 1977; Fenton et al., 1977) might therefore be caused mainly by the disruption of this exosite due to excision of peptide Ile 68–Arg 77A (Hofsteenge et al., 1988), i.e., similar to  $\beta$ -thrombin.

Peptide segment Thr 147–Ser 153 might be involved in thrombomodulin binding (Suzuki et al., 1990). Other evidence (Hofsteenge et al., 1986; Preissner et al., 1987; Noé et al., 1988; Jakubowski & Owen, 1989; Tsiang et al., 1990) indicates, however, that the accessibility and integrity of the 70–80 surface loop and the fibrin(ogen) secondary binding exosite are at least also required for efficient thrombin interaction with thrombomodulin. It has been shown (LeBonniec & Esmon, 1991) that a Glu 192  $\rightarrow$  Gln substitution in thrombin mimics the catalytic switch induced by thrombomodulin. Thus, the thrombin interaction site for this endothelial cell-surface protein seems to comprise a larger surface area including the junction between loop segments 70–80 and 144–151 (Wu et al., 1991).

#### *The putative heparin-binding site*

A thrombin surface patch of positively charged side chains (Arg 126, Lys 236, Lys 240, and Arg 93) surrounded by other basic residues (Arg 101, Arg 233, Arg 165, Lys 169, Lys 235, Arg 175, Arg 173, Arg 97, patch p2 in Table 7), which in turn contact negatively charged side chains (see also Tables 5, 6), is shown in Figure 18A. These positively charged residues are quite densely packed on the surface and are not compensated by negatively charged groups below the surface as for the 70–80 loop. Each of them therefore gives rise to a very strong positive electrostatic field (Fig. 8). In spite of this strong field, we see no evidence for fixed solvent anions (Fig. 18B); the isolated blobs of electron density within the surface depression and in the intermolecular (crystal) gap are sufficiently interpreted as water molecules. This site seems rather to be more suited to binding polyanions such as

heparin; this glycosaminoglycan of  $1\frac{1}{2}$  acidic (sulfate and uronic) groups per monosaccharide unit binds to thrombin (Griffith et al., 1979) and accelerates its inactivation by antithrombin III (Björk & Lindahl, 1982), presumably through interactions with alkaline groups of thrombin and antithrombin.

Chemical modification experiments on thrombin revealed that Lys 169 and Lys 240 (at the edges of this positively charged patch [Fig. 18A]) are blocked by bound heparin and that their modification renders thrombin inaccessible to heparin (Church et al., 1989). Cooperative electrostatic interactions of linear heparin molecules (with high linear negative charge density) with this positively charged region of contiguous basic residues could explain the high affinity for binding heparin to thrombin (Heuck et al., 1985). Furthermore, this putative heparin-binding site exhibits a typical heparin-binding motif (a helical strand comprising several basic residues [see Table 3; Cardin & Weintraub, 1989]). Some slippage of this thrombin helix caused by heparin binding cannot be excluded bearing in mind the relatively variable arrangement of the carboxy-terminal helix in related serine proteinase-reactive domains (Bode & Huber, 1986).

Acceleration of complex formation between antithrombin III and thrombin by heparin may be due in part to a template mechanism (Pomeranz & Owen, 1978; Griffith, 1982; Nesheim, 1983) in addition to its effects on the antithrombin III structure. If a docking geometry to be similar to that known for the "smaller" protein inhibitors (Bode & Huber, 1991) is assumed for antithrombin III bound to thrombin, the thrombin heparin-binding site would be located near the presumed heparin-binding site of antithrombin III in antithrombin III–thrombin complexes (for references, see Huber & Carrell, 1989). It is therefore conceivable that one heparin strand of about 14 monosaccharide units (required for efficient acceleration of the thrombin–antithrombin III reaction [see Huber & Carrell, 1989]) could connect the two components. The Arg 97 and Arg 175 residues, which form a positively charged satellite patch (p2c in Table 7; Figs. 7, 8, 18) to the "north rim" (Bode et al., 1989b) of the active-site canyon could further stabilize such a mediatory binding through heparin.

Similar reaction rates of antithrombin III with  $\alpha$ - and with  $\gamma$ -thrombin in the presence of heparin support the idea that the main heparin recognition sites of antithrombin III-bound thrombin reside outside the fibrin(ogen) secondary binding exosite (Stone & Hofsteenge, 1987). Other evidence (Olson et al., 1986; Hogg & Jackson, 1989; Naski et al., 1990) indicates, however, that heparin might bind to both exosites in the absence of antithrombin III. This highly charged putative heparin-binding site might also (in cooperation with the fibrin(ogen) secondary binding exosite) mediate thrombin binding to cation exchange resins and fibrin (Fenton et al., 1988) and might therefore be alternatively called anion-binding exosite II.



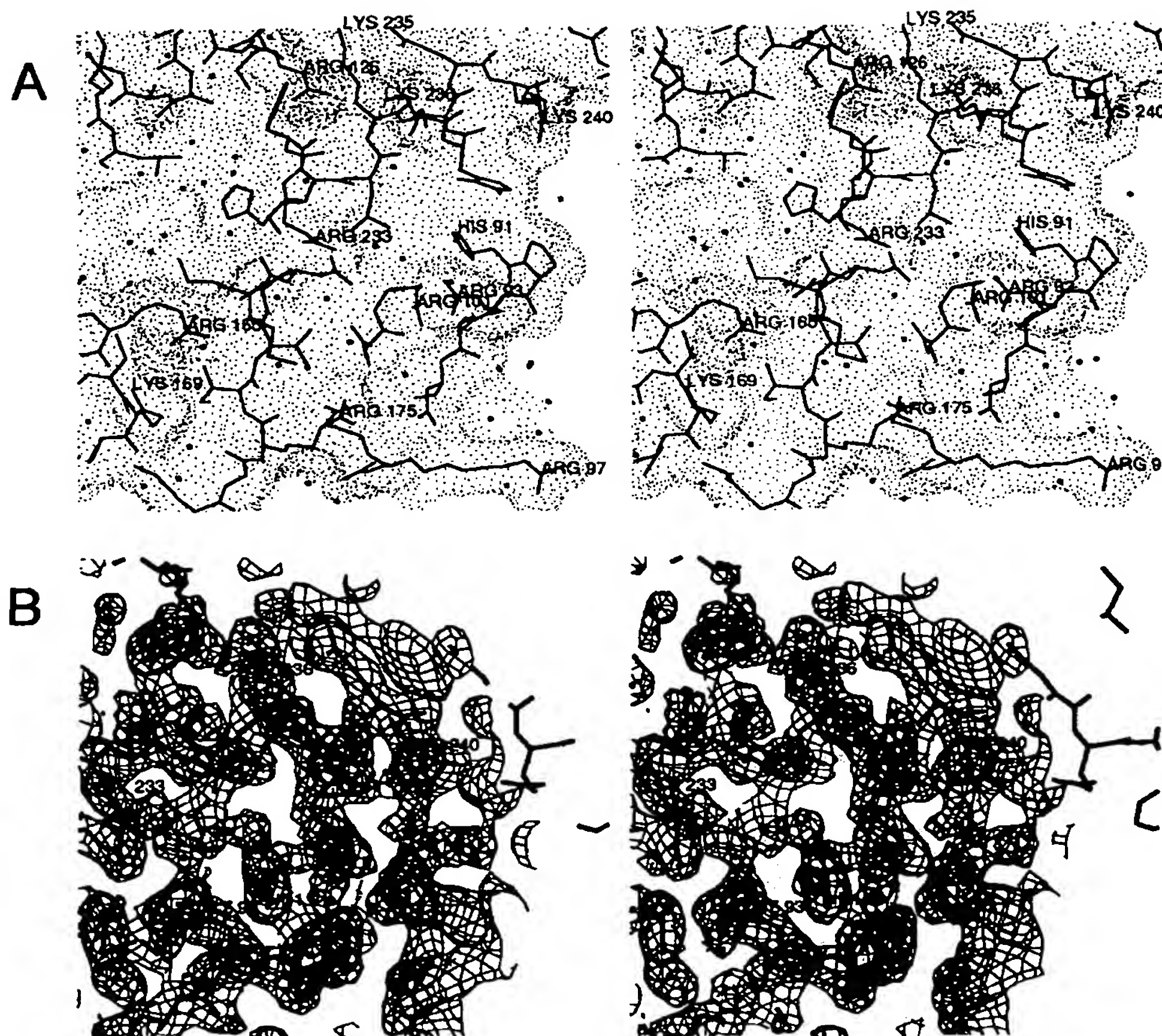


Fig. 18. Putative heparin-binding site of human  $\alpha$ -thrombin. The view is obtained from the standard view upon an approximately  $90^\circ$  rotation around a horizontal axis. A: Surface structure between Lys 169 (lower left corner) and Lys 240 (upper right corner), displayed together with the Connolly surface. The site of highest positive electrostatic potential is in the surface depression close to His 91, surrounded by Arg 126, Lys 236, Lys 240, Arg 93, Arg 101, and Arg 233. Crosses represent localized solvent molecules. B: Section of the final electron density superimposed with equivalent surface parts of thrombin and a few localized solvent molecules (crosses). The depression outside the thrombin surface does not contain appropriate high density, which could account for a bound anion. Contour surface at  $0.8\sigma$ .

Furthermore, the heparin-binding site is close to the amino-terminal Thr 1H of the A-chain (see Fig. 7); its partial shielding by adjacent pro-parts in prothrombin (Fenton et al., 1988) or in other partially activated prothrombin forms still possessing pro-parts (such as meizothrombin [see Doyle & Mann, 1990]) also seems possible therefore without conformational changes in the B-chain globule. Chemical modification experiments further suggest that part of this heparin-binding site (in particular Lys 169 and/or Lys 240) might act simultaneously as a high-affinity binding site for platelets (White et al., 1981).

In the natural human variant thrombin Tokushima, Arg 101 is replaced by a tryptophan residue (Miyata et al., 1987). The side chain of Arg 101 points away from the active-site cleft (Fig. 12) but is involved with the adjacent

Asp 100 in a surface-located salt bridge (Table 6). According to modeling experiments, substitution with a tryptophan side chain might disturb the "rim" region and would thus primarily affect binding of more extended substrates and interactions with cells. An effect on the subsequent active-site residue, Asp 102, whose side chain points in the opposite direction (toward the active-site cleft [see Fig. 1]), cannot be excluded; it would be small, however.

#### *The thrombin RGD site*

Both human and bovine thrombin contain an RGD segment (Arg 187–Gly 188–Asp 189) beneath the specificity pocket (Fig. 19). This segment has recently been suggested (Fenton, 1988) to play a distinct role as a docking

site in platelet binding (e.g., through the platelet-specific GPIIb/IIIa integrin receptor [see Charo et al., 1987]). Glenn et al. (1988) showed that a thrombin-derived 23-amino acid peptide comprising this Arg-Gly-Asp region competes with  $\alpha$ -thrombin for binding to the high-affinity receptor on fibroblasts, and that it favors the thrombin-mediated thymidine incorporation; in contrast, the RGDA peptide, which likewise competes with  $\alpha$ -thrombin, inhibits the mitogenic effect of  $\alpha$ -thrombin. Bar-Shavit et al. (1991) demonstrated recently that the endothelial cell adhesion properties of modified human thrombin could be inhibited specifically by a GRGDSP peptide; antithrombin III inhibited this thrombin adhesion, whereas hirudin did not interfere with it.

In the PPACK- $\alpha$ -thrombin structure, the side chain of Arg 187 runs along the molecular surface; it is exposed to solvent with the distal guanidyl group engaged in quite strong salt/hydrogen-bond bridges (see Fig. 19; Table 6). Neither Gly 188 nor Asp 189 (except for its carboxylate group, which points toward the interior of the specificity pocket [Fig. 19]) are, however, accessible to receptor structures of approaching cells or proteins without complete unfolding of this thrombin site. Such an accessibility would seem to be required, however, for proper recognition of the underlying sequence motif. In partially degraded or thermally unfolded thrombin derivatives or in zymogenlike (pro)thrombin species (with the Asp 189-containing peptide fragment turned inside-out as in chymotrypsinogen, see above) a greater exposure of this RGD site is conceivable; this would be in agreement with the observation that intact  $\alpha$ -thrombin attains endothelial cell adhesive properties only after exposure to elevated temperatures (Bar-Shavit et al., 1991).

### The active-site cleft

The most remarkable feature of the thrombin surface is a deep narrow "canyon" (see Fig. 1) containing the catalytic residues and the adjacent substrate-binding subsites (Bode et al., 1989b). The catalytic residues (in particular Ser 195, His 57, and Asp 102) divide this cleft into two halves, corresponding to a fibrinopeptide side ("west") and a fibrin side ("east"). This cleft is bordered and shaped by the four extending polypeptide segments, Arg 173-Ile 174, Asp 95-Leu 99, Tyr 60A-Phe 60H, and Phe 34-Cys 42, on one side (forming the "north rim," Fig. 1) and the two loop segments, Gly 216-Cys 220 and Asn 143-Gln 151, on the opposite side (the "south rim"). The side chains of Ile 174, Tyr 60A, Trp 60D, Lys 60F, Phe 60H, and of Thr 147, Trp 148, respectively, protrude into the active-site cleft, i.e., each rim is lined primarily by hydrophobic groups. The base, in contrast, is covered mainly with polar/charged side chains (Glu 217, Glu 192, Asn 143, Gln 151, Arg 73) and polar main-chain groups.

Around the entrance to the specificity pocket (Fig. 20), 10 negatively charged residues (among them Asp 189, Glu 192, Asp 221, and Glu 146) are clustered (patch n1a in Table 7), with Glu 192 (relatively flexible at the base) and Asp 189 (at the bottom of the specificity pocket [see Fig. 20]) being the only acidic residues not involved in direct salt bridges. The latter two, together with Asp 221, Glu 18, Asp 186A, and Glu 186B, exert repulsive forces on one another and contribute substantially to the rather negative field generated at this site (see Fig. 8; Tables 5, 7). According to our calculations, the catalytic residues are only marginally affected by this negative electric potential (due to their position at the edge of the negative

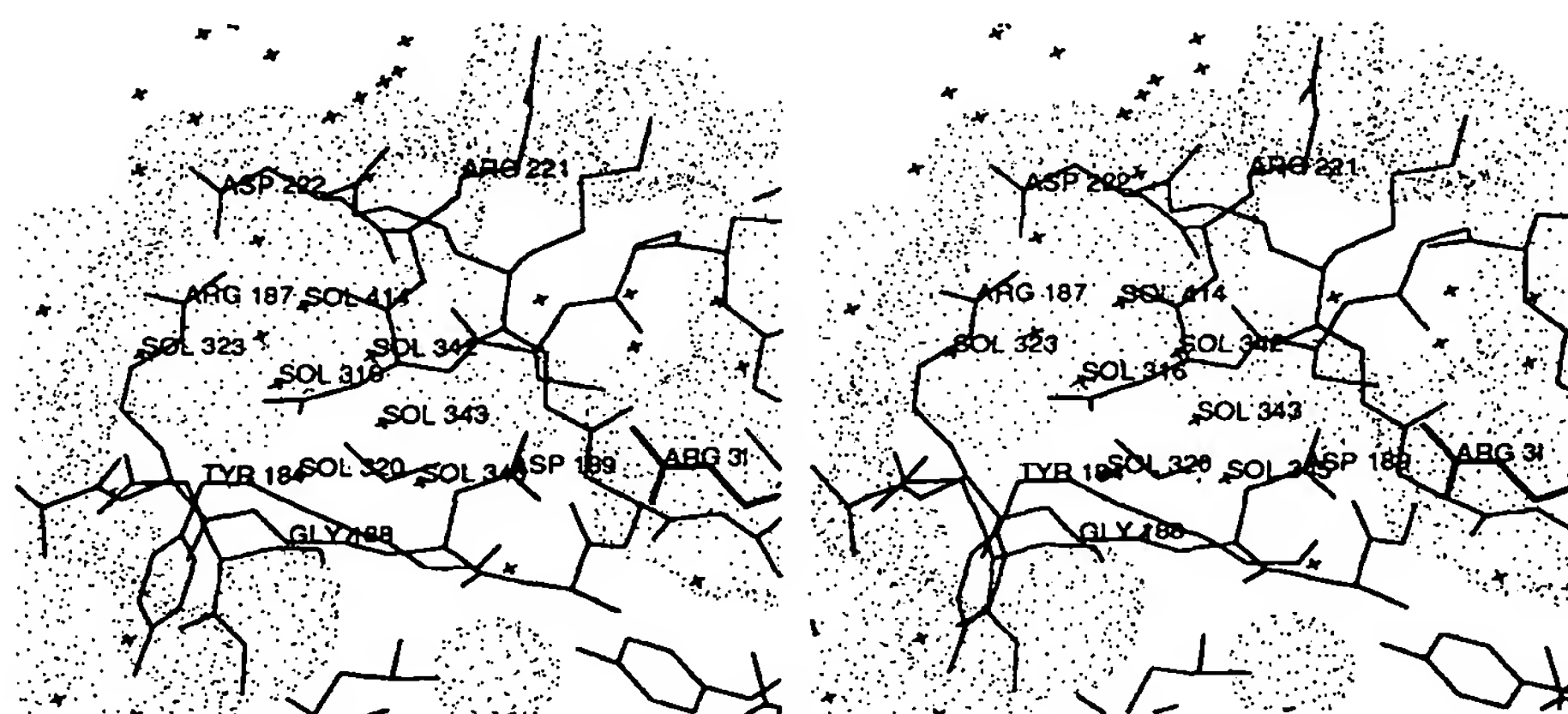


Fig. 19. Segment Arg 187-Gly 188-Asp 189 and adjacent residues of human  $\alpha$ -thrombin (thin connections), displayed together with the Connolly surface. Most of the internal solvents (crosses), which form an extended solvent cluster, are labeled. Arg 187 is located at the thrombin surface; Gly 188 and Asp 189 are, however, buried in the interior of the thrombin molecule. The carboxylate group of Asp 189 is directed toward the specificity pocket and in ion pair contact with the guanidyl group of Arg 31 of PPACK (thick connections). The main-chain angles of this segment are  $-138^\circ$ ,  $174^\circ$  (Arg 187);  $132^\circ$ ,  $155^\circ$  (Gly 188);  $-171^\circ$ ,  $16^\circ$  (Asp 189). The view is similar to the standard view of Figure 1.



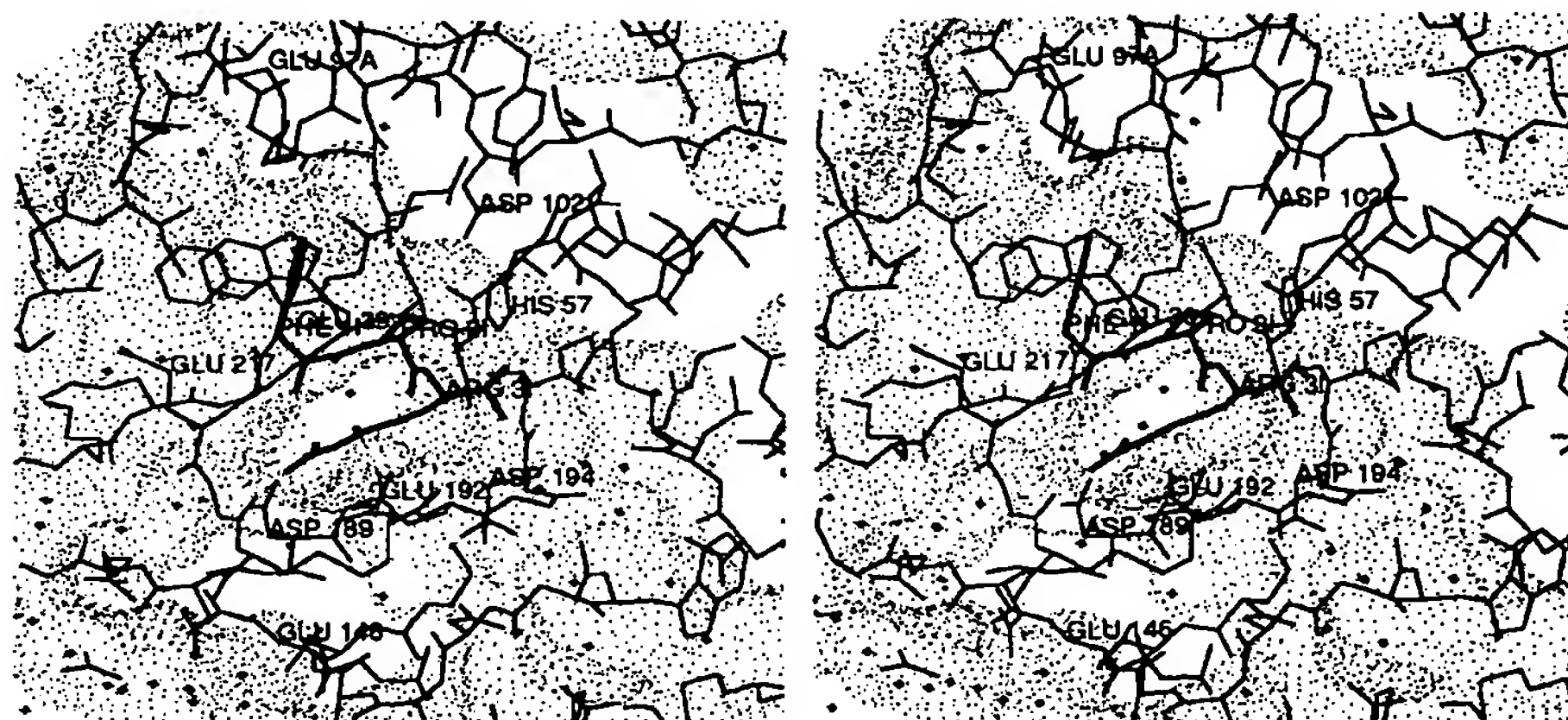


Fig. 20. The active-site region of PPACK-thrombin, displayed together with the Connolly surface of thrombin. The thrombin structure is given with thin connections, the PPACK molecule with thick connections, and localized solvent molecules are represented by crosses. The 60 insertion loop, which partially occludes the active-site residues and the entrance to the specificity pocket, is cut off; thus, the active-site residues and Pro 21 and Arg 31 are freely visible. The negatively charged residues arranged around the specificity pocket to contribute to the negative patch n1a (Table 7) are labeled. Standard view as in Figure 1.

charge-patch); besides their mutual interactions, they experience mainly the ammonium group of Lys 60F. The latter interaction should shift the pK values of Asp 102 and His 57 to slightly lower values than normal. As described above, the fibrin(ogen) secondary binding exosite (at the right-hand side of the active-site cleft [Figs. 1, 14]) in contrast gives rise to a strong positive electrostatic field (see Fig. 8). This dipolar charge distribution along the active-site cleft would clearly influence the orientation of approaching substrates and inhibitors of large electric moments (such as hirudin, see above).

#### *The thrombin active site and interaction with PPACK*

The PPACK molecule, as it is bound to the active-site cleft of thrombin, is shown in Figure 21 and Kineimage 6. Its peptidyl moiety juxtaposes the extended thrombin segment Ser 214–Glu 217 in a twisted antiparallel manner. The carboxy-terminal Arg 31 carbonyl group is part of a tetrahedral hemiketal structure; i.e., it is linked covalently to Ser 195 O $\gamma$  and, via the methylene group, to the imidazole group (N $\epsilon$ 2) of His 57 of throm-

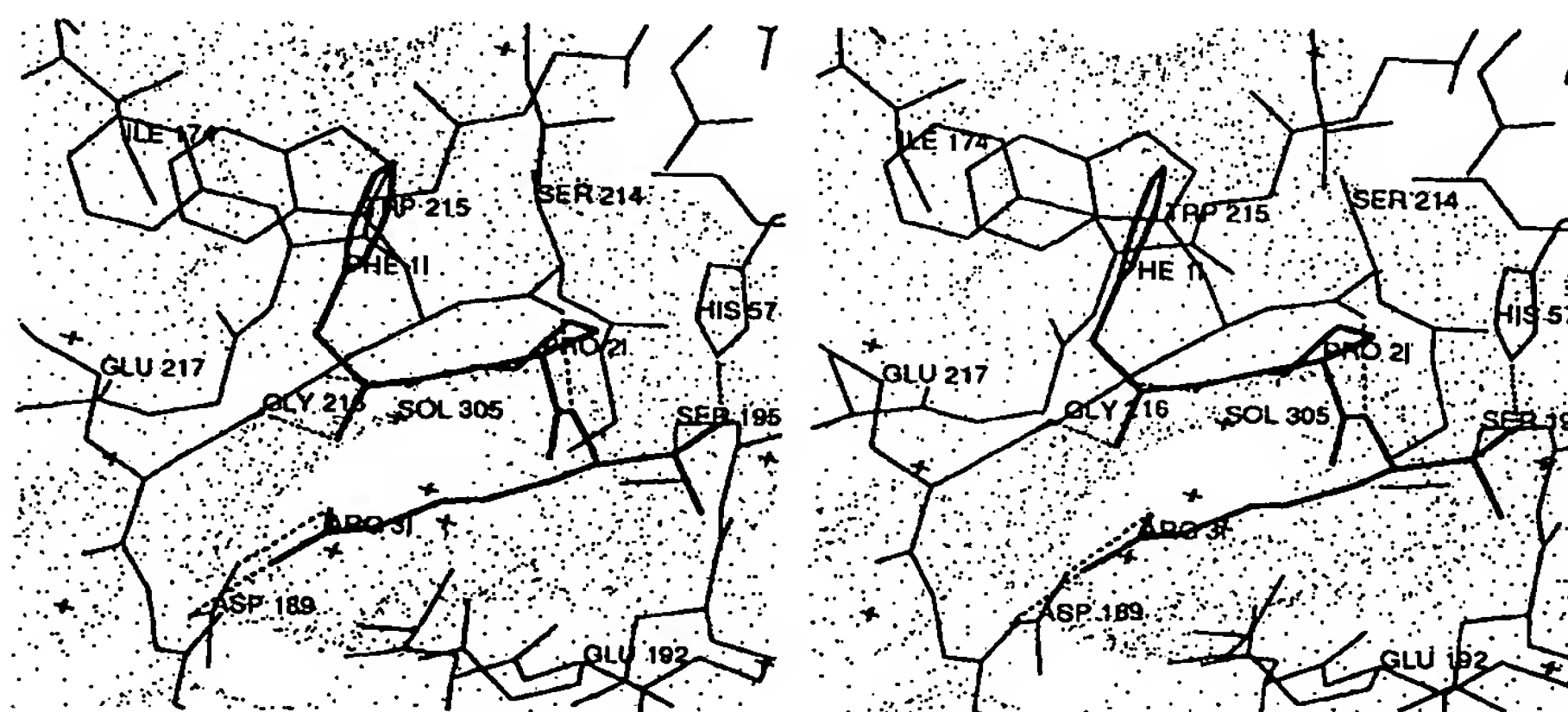


Fig. 21. Interaction of the PPACK molecule (thick connections) with the thrombin active site (thin connections), displayed together with the Connolly surface of thrombin and some localized solvent molecules (crosses). Most of the hydrogen bonds formed between PPACK and thrombin are shown in addition (dashed lines). The Arg 31 carbonyl group forms a tetrahedral hemiketal structure covalently linked with Ser 195 O $\gamma$  and (via the methylene group) with His 57 N $\epsilon$ 2. Standard view as in Figure 1.

bin. The side chain of Ser 195 is probably considerably rotated to form this covalent bond compared to its location in unliganded thrombin.

The thrombin specificity pocket is geometrically quite similar to that of bovine trypsin (Fig. 22) (Bode & Schwager, 1975a). The r.m.s. deviation of 22  $\alpha$ -carbon atoms of the residues and segments lining the specificity pocket in PPACK-thrombin is 0.39 Å compared with bovine trypsin; an almost identical deviation is obtained upon comparison of PPACK-thrombin with hirudin-thrombin. The only thrombin residue whose side chain points into the pocket differently is Ala 190, replacing a serine residue in trypsin. This exchange renders the thrombin pocket slightly less polar but, in particular, causes the loss of a hydrogen-bond acceptor/donor (perpendicular to the guanidyl group of Arg 31 [Fig. 22]). Solvent molecule 305, which becomes buried upon complex formation, is conserved in the thrombin pocket (compare with trypsin [Bode & Schwager, 1975a]). This solvent molecule finds a much more hydrophobic surrounding in thrombin than in trypsin, but is nevertheless adequately positioned to receive a favorable hydrogen bond from the NH1 atom of the Arg 31 guanidinium group. In consequence, the thrombin S1 subsite would better accommodate an arginine P1 side chain, whereas an inserted lysine side chain would fit less tightly in the thrombin pocket, due to the greater space and lack of a hydrogen-bond partner (residue 190). This is in agreement with experimental results concerning the relative affinities for arginine and lysine P1 side chains (Liem & Scheraga, 1974; Kettner & Shaw, 1979, 1981; Lottenberg et al., 1983).

The amidino nitrogen of the P3 residue Arg 31 is 2.9 Å away from Ser 214 O, and in a relatively favorable orientation to form a hydrogen bond (Fig. 21). Formation of this hydrogen bond is assumed to be an important step to achieve the tetrahedral transition state in productive enzyme-substrate complexes; thus, the stable state seen here mimics somewhat a real transition state.

Quick II (Henriksen & Mann, 1989), another naturally occurring dysfunctional thrombin mutant, has the replacement Gly 226  $\rightarrow$  Val 226 (see Fig. 12). Modeling ex-

periments show that the bulky valine side chain would probably disturb the Asp 189 conformation at the bottom of the pocket, rather than directly blocking binding of an inserted arginine side chain. The pocket of Quick II will probably not be as narrowed as in porcine pancreatic elastase (Meyer et al., 1988) or human leukocyte elastase (Bode et al., 1986). Thus, neither typical trypsin nor elastase substrates should be cleaved efficiently by this mutant, in agreement with the experimentally observed lack of catalytic and clotting activity (Henriksen & Mann, 1989).

Behind Asp 189, the thrombin insertion loop Tyr 184A-Gly 188 and segment Asp 221A-Tyr 225 enclose a large internal cavity filled with (buried) solvent molecules (see Fig. 19; Table 9). This thrombin water cluster is even larger than in trypsin, due to the smaller Thr 172 of thrombin compared with a tyrosine in trypsin. This cluster and its linkage to the bulk water have previously been suggested as providing a route to expel water molecules through the back of the specificity pocket upon insertion of a P1 side chain of a binding peptide (Meyer et al., 1988).

Another more extended, quite hydrophobic pocket lies near the entrance to the specificity pocket (Fig. 21), lined by Ile 174, Trp 215, segment 97-99, His 57, Tyr 60A, and Trp 60D. This hydrophobic depression is most certainly the suggested apolar binding site (Berliner & Shen, 1977). Its back part is made up of the same structural elements as in trypsin (Trp 215, a shorter segment 97-99, His 57 [see Fig. 22]); it is in thrombin, however, further screened off from bulk water by the mainly hydrophobic Tyr 60A-Trp 60D loop on one side and Ile 174 on the other side (to the right and left in Fig. 21, respectively). The PPACK molecule nestles tightly into this depression. The solvent-accessible surface of the PPACK-thrombin complex is 240 Å<sup>2</sup> smaller than that of thrombin itself; only one-third of the PPACK surface (one edge of the benzene ring of D-Phe 11, and Pro 21 [see Fig. 21]) does not become buried upon the complex formation. In the PPACK-thrombin complex the pyrrolidine ring of Pro 21 is almost fully shielded from bulk water; it is in a completely hy-

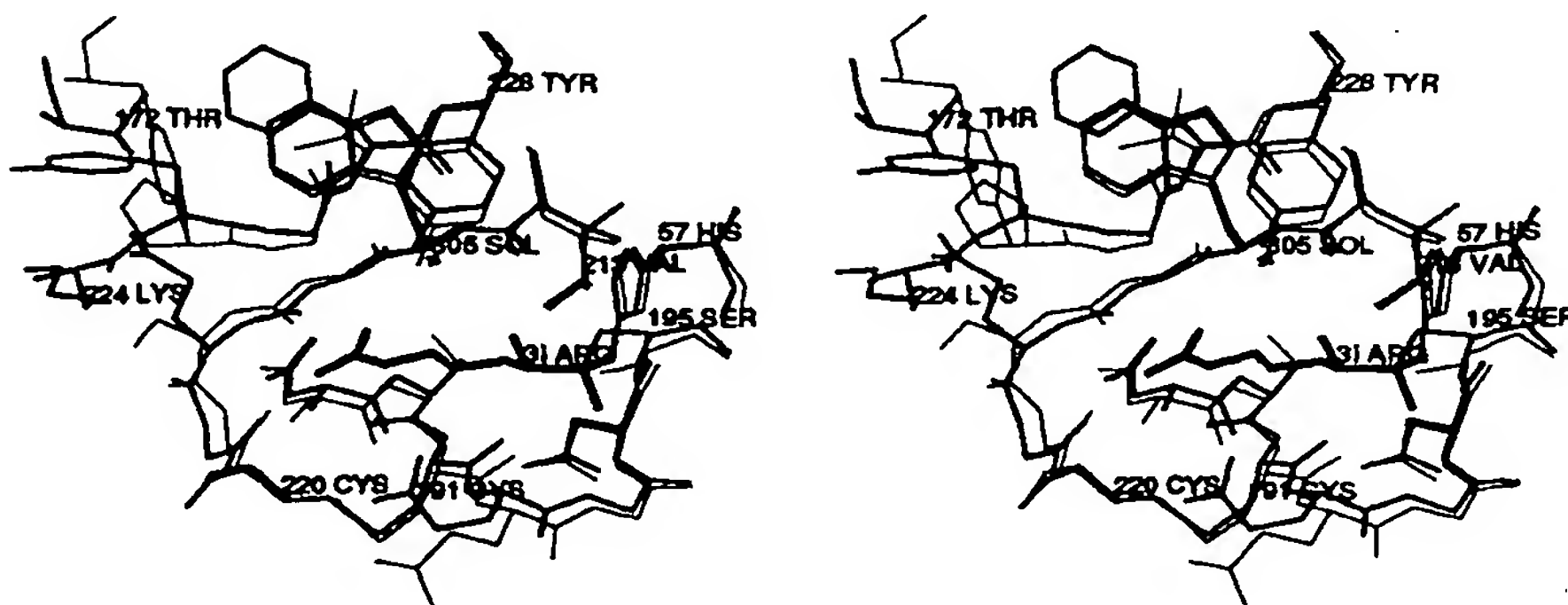


Fig. 22. Human  $\alpha$ -thrombin structure around the specificity pocket (medium connections) superimposed by the equivalent residues of unliganded bovine trypsin (thin connections [Bode et al., 1976]). Arg 31 of PPACK is shown with thick connections. The view is into the specificity pocket, i.e., similar to the standard view of Figure 1. Ser 190 of trypsin is replaced in thrombin by an alanine residue. Sol 305 in thrombin has an equivalent water molecule (Sol 416) in trypsin.

drophobic environment and seems to be perfectly accommodated through its polyproline II conformation (see Fig. 21). This excellent packing in the thrombin S2 cavity explains the frequent occurrence of proline (and other medium-sized hydrophobic) residues at the P2 position of naturally occurring thrombin macromolecular substrates. In contrast to trypsin (Liem & Scheraga, 1974; Bajusz et al., 1978), it further accounts in thrombin for the particular reactivity of chloromethylketone inhibitors with P2-proline residues (Kettner & Shaw, 1981). Much larger residues would not pack favorably into the S2-thrombin cavity, and smaller ones (such as glycine residues) would not be as well fixed.

The (substituted) piperidine rings of the benzamidine (NAPAP [Stürzebecher et al., 1983]) and of the arginine-based (MQPA [Okamoto et al., 1981]) synthetic inhibitors (Fig. 23) occupy this pocket where they can be tightly fixed and sandwiched between the flat side of the His 57 imidazole ring and the amino-terminal aryl moiety (Bode et al., 1990; Turk et al., 1991; Brandstetter et al., 1992). As in the case of PPACK-thrombin, the hydrogen bonds made with Gly 216 N and O determine mainly the docking position of these inhibitors in the thrombin active site.

The side chain of the preceding residue of PPACK, D-Phe 11, fits into the notched, mainly hydrophobic cleft made by Ile 174, Trp 215, segment 97–99, and Tyr 60H (Fig. 21). Such an interaction is not only preferred due to favorable hydrophobic contacts; presumably, both the perpendicular aryl-aryl (edge-on) arrangement (Burley & Petsko, 1985) and the aryl-carbonyl contact with the carbonyl group of Glu 97A (see Thomas et al., 1982) contribute significantly to binding strength (Bode et al., 1990). In natural protein substrates, the variety of amino

acids in P3 is great; in general, L-amino acids in P3 position would interact with a different thrombin site. The P4 side chain of a bound all-L-amino acid polypeptide substrate would, however, point toward this hydrophobic cleft; in fact, the beneficial effect of large bulky aromatic residues at P4 connected to a P3 glycine has been used to construct suitable thrombin substrates (Svendsen et al., 1972). Very recently we have shown that the naphthyl and tosyl moieties of certain benzamidine-derived synthetic inhibitors (see Fig. 23; Bode et al., 1990; Turk et al., 1991) as well as other aromatic groups in arginine-derived synthetic inhibitors (Matsuzaki et al., 1989) interact in a similar favorable manner with this cleft (Brandstetter et al., 1992). The phenolic side chain of Tyr 3HI of hirudin is also placed in such a manner (Rydel et al., 1991), and there is now crystallographic evidence (Martin et al., 1992; Stubbs et al., 1992) for an equivalent arrangement of the phenyl moiety of the fibrinogen A $\alpha$ -chain Phe 8 residue upon binding to thrombin similar to previous suggestions (Meinwald et al., 1980; Ni et al., 1989a,b). We have therefore proposed the designation aryl-binding site (Bode et al., 1990), distinguishing it from the other part of the apolar binding site (the S2 cavity).

The extraordinary specificity and reactivity of PPACK for thrombin (Kettner & Shaw, 1979) can thus easily be attributed to the optimal fit of all three residues to their respective thrombin subsites (see Fig. 21). In addition, each of these PPACK residues seems to exist in a stereochemically and energetically optimal conformation in the thrombin-bound PPACK molecule (see Table 10): both the main- and the side-chain moieties of Arg 31 are arranged in an extended conformation of favorable energy; Pro 21 exhibits a characteristic polyproline II conforma-

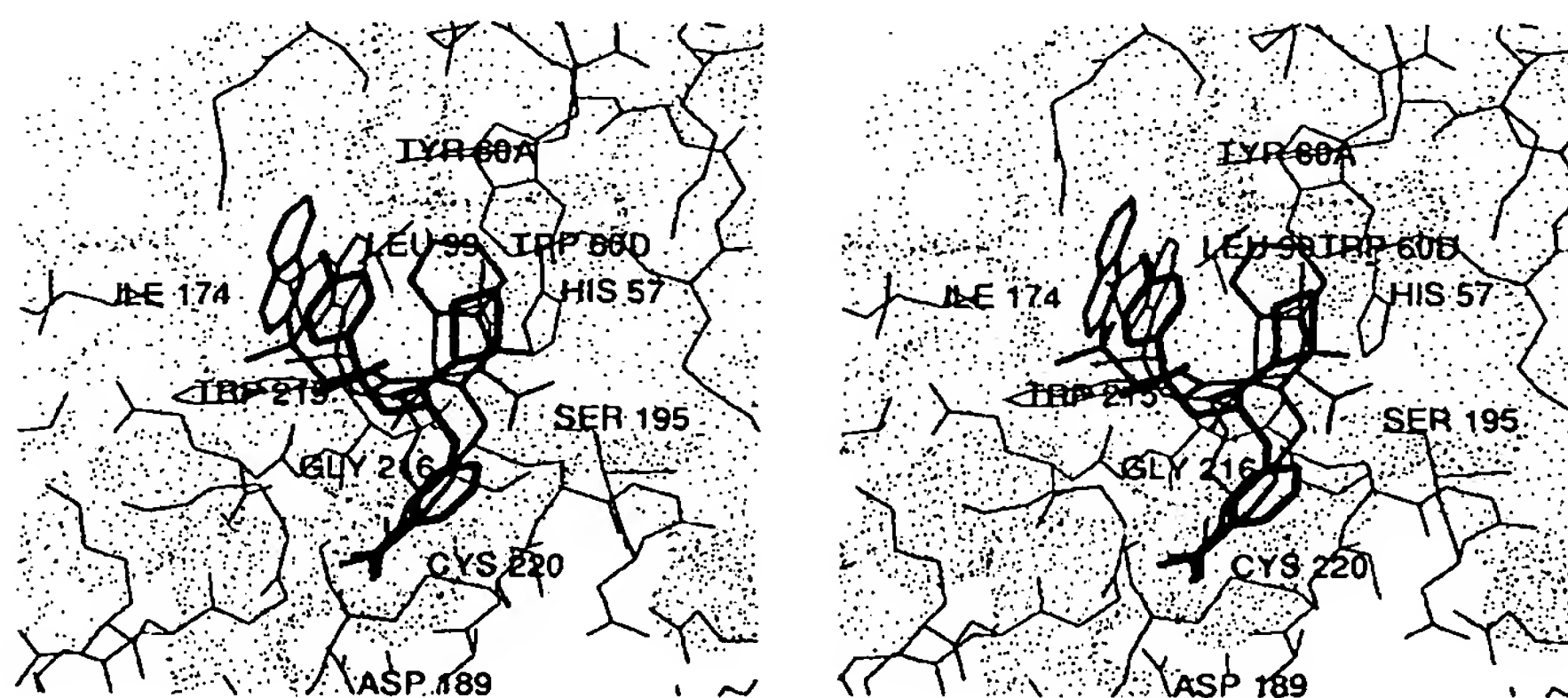


Fig. 23. Tosyl-*m*-amidinophenylalanyl-piperidine (thick connections), NAPAP (medium connections), and MQPA (thin connections) bound to the active site of human  $\alpha$ -thrombin displayed together with the Connolly surface of thrombin (Turk et al., 1991). The naphthyl/toluene/chinolyl groups of the inhibitors interact with the aryl-binding site of thrombin; the side chains of the *m*- and the *p*-amidinophenylalanyl residues and of the arginyl residue enter the specificity pocket from slightly differing sites; the S2 subsite of thrombin is occupied to different extents by the (partially substituted) piperidine moieties. The view is similar to the standard view of Figure 1.



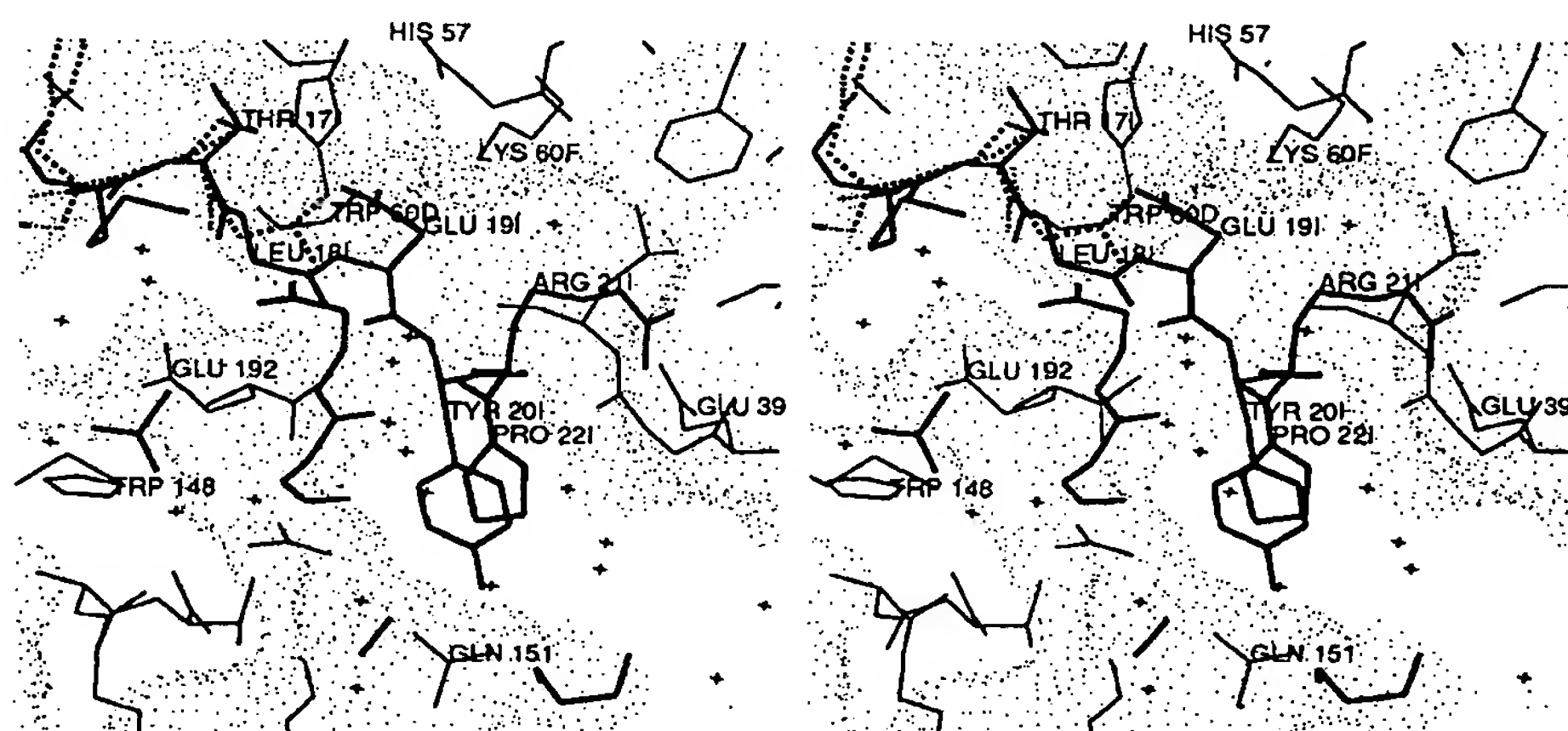
tion; and D-Phe 11 (with a  $C\alpha$ -C angle of  $-131^\circ$ ) represents just one of the two optimal conformers preferred for steric reasons by a D-amino acid residue linked to a proline.

The ammonium and the carbonyl group of D-Phe 11 are in favorable hydrogen-bond contacts with Gly 216 (2.7 Å [see Fig. 21]). In contrast to trypsin, good peptidyl aldehyde inhibitors of thrombin require termination at P3 (Bajusz et al., 1978), probably to allow the undisturbed formation of this hydrogen bond; it is clear, however, that *N*-alkylation of PPACK (such as in *N*-methyl-PPACK [see Bajusz et al., 1990]) should not impair hydrogen-bond formation in solution. In the PPACK-thrombin crystal structure, the ammonium group forms two further hydrogen bonds with a fixed solvent molecule and with Gln 244 of a symmetry-related thrombin molecule. This contact might confer stability to this PPACK-thrombin crystal form.

An L-configured Phe 11 residue at P3 would not be able to interact with thrombin in as favorable a manner as the D-diastereomer (in particular when linked to a proline [see Fig. 21]). Simultaneous hydrogen-bond formation through its amino group and favorable aryl interaction are mutually exclusive; the best fit for an L-Phe 11 residue would seem to be at a  $C\alpha$ -C dihedral angle of about  $-60^\circ$ , with the phenyl group fitting into the aryl-binding site in a slightly different relative orientation.

Due to lack of direct structural evidence, the primed subsites of thrombin cannot be described with a confidence equivalent to the nonprimed sites. The thrombin

complexes of various protein inhibitors (see, e.g., Fig. 24), modeled according to experimental serine proteinase complexes, might give a rough idea of the probable interactions of substrate P1', P2', and P3' residues with thrombin. The S1' subsite of thrombin is (due to the impending Lys 60F) limited in size and therefore particularly suited for small P1' residues such as glycine or alanine (or even serine or aspartate); this would be in agreement with the frequent occurrence of small P1' residues in natural protein substrates (see, e.g., Lottenberg et al., 1983) and with thrombin's preference for alanine and glycine over leucine at this substrate position (Liem & Scheraga, 1974). P2' side chains of protein inhibitors would normally interact with a surface provided by the side chain of residue 151 of the cognate serine proteinase (see Fig. 24). A proline residue is present at the P2' position of the fibrinogen A $\alpha$ -chain, and synthetic peptides with a proline at P2' have been found to be readily cleaved by thrombin (Liem & Scheraga, 1974); in light of the presumed importance of the hydrogen bond formed between the P2' amide nitrogen and the carbonyl group of residue 41 to achieve the transition state in productive enzyme-substrate complexes (Bode & Huber, 1986; Read & James, 1986), this preference for proline residues (which cannot mediate such a hydrogen bond) is remarkable. Inspection of the structure shows that this P2' proline could be accommodated in a hydrophobic pocket (left to Leu 41 [Fig. 24]) with the following residues of a thrombin-bound peptide probably taking a different route than the binding loop of canonical protein inhibitors (Stubbs



**Fig. 24.** Interface of the hypothetical complex of human thrombin with a Kazal-type inhibitor. The primed subsites of human thrombin (thin connections) are displayed together with the Connolly surface of thrombin; superimposed are the PPACK molecule (dashed connections) and part of the turkey ovomucoid third domain (thick connections) as determined in its complex with human leukocyte elastase (Bode et al., 1986) after fit of the elastase component to thrombin. The inhibitor scissile peptide bond is between Leu 181 (P1) and Glu 191 (P1'). The side chain of Glu 191 would collide with the indole ring of Trp 60D of the (partially cut-off) 60 insertion loop; only a shorter side chain should become accommodated at S1'. There is space for a Tyr 201 (P2') side chain; the interactions with thrombin would, however, not be optimal. The side chain of Arg 211 (P3') would collide with Glu 39; however, other more favorable arrangements can be assumed. The view is similar to the standard view of Figure 1.



et al., 1992). The relatively beneficial effect of an arginine residue at P3' (Liem & Scheraga, 1974) could then be attributed to ionic interactions with glutamic acid residues 192 and 39 of thrombin (Fig. 24), in agreement with recent results obtained with a Glu 39 → Arg mutant (Le-Bonniec et al., 1991), and that of a valine residue at P4' by hydrophobic interactions with Phe 60H.

*Probable interaction with macromolecular substrates and inhibitors*

The narrow, canyonlike structure of the thrombin active-site cleft is clearly the main source of the remarkable thrombin specificity toward macromolecular substrates and inhibitors (Bode et al., 1989b). It restricts access to the thrombin catalytic center of potential scissile peptide bonds of most macromolecules simply through steric hindrance, thereby discriminating in this way susceptible substrates by means of shape and conformation. An attempt to dock the archetypal inhibitor bovine pancreatic trypsin inhibitor (BPTI) to the thrombin active site with identical position and orientation as observed for the trypsin(ogen) complexes (Fig. 25) (Huber et al., 1974; Bode et al., 1978, 1984; Huber & Bode, 1978) results in a collision of residues of the inhibitor-binding region (in particular Tyr 35I and segment 37I–40I) with part of the 60 insertion loop of thrombin (in particular with the side chain of Trp 60D [see Fig. 25]).

It appears as if this steric hindrance cannot be relieved, even upon considerable relative tilting of the inhibitor. The inherently flexible Trp 148 loop (left-hand side in Fig. 25) could give way; the 60 insertion loop (right-hand side, Fig. 25), however, appears so rigid that strong forces would be required to allow access of the inhibitor. These docking experiments agree with experimental results, which reveal extremely low association constants of about

$10^3 \text{ M}^{-1}$  for  $\alpha$ -thrombin, but almost 10 times larger values for the more open and flexible  $\gamma$ -thrombin (Berliner et al., 1986; Ascenzi et al., 1988). The structure of the complexes formed at very high BPTI concentrations is unclear; however, an enforced expulsion of the 60 insertion loop by BPTI (in a similar manner to when the non-complementary binding site of trypsinogen is forced into a fitting trypsinlike structure upon BPTI complexation [see Bode et al., 1978; Huber & Bode, 1978; Bode, 1979]) and consequent binding at the expense of free energy cannot be ruled out.

Similar docking experiments with other "small" serine proteinase protein inhibitors of known spatial structure such as eglin c (Bode et al., 1987), the squash seed inhibitor CMTI-I (Bode et al., 1989a), and the turkey ovomucoid third domain (Bode et al., 1986) show similar restraints (see Fig. 24). In each case the binding loop residues on both sides of the scissile peptide bond would essentially fit to the corresponding subsites around the thrombin catalytic site; however, collision with the thrombin "rims" would occur through structural elements extending from the inhibitor core. Of the protein inhibitors tested for regard to possible thrombin inhibition so far, the Kazal-type inhibitors seem to cause least steric hindrance (in spite of the noninhibiting behavior reported for ovomucoid [Berliner et al., 1986]); they therefore seem to be the most promising candidates for future protein design experiments.

In the absence of an experimental structure of an intact functional serpin inhibitor or of a serpin-serine proteinase complex, one can only speculate on the interaction mode between thrombin and antithrombin III. There are, however, several lines of evidence (see Bode & Huber, 1991) that serpins bind in an essentially similar manner as the "small" protein inhibitors to their cognate enzymes,

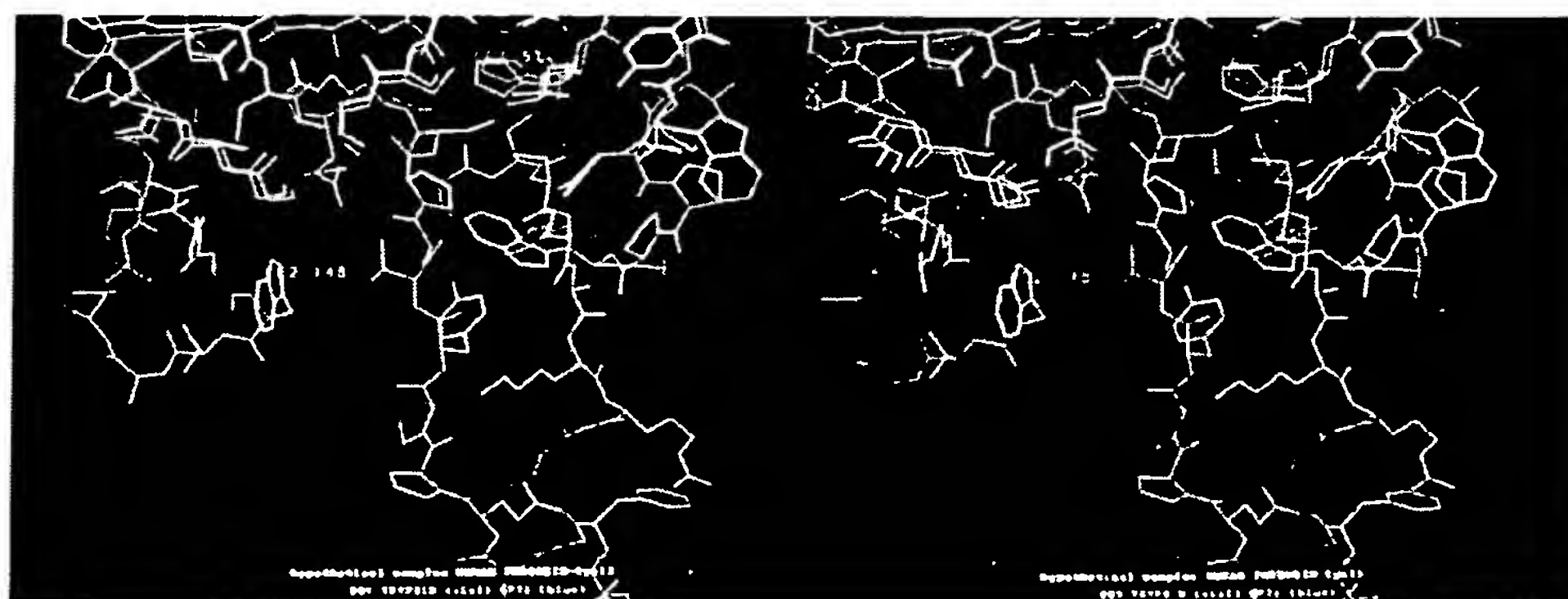


Fig. 25. Part of the hypothetical docking complex formed between human  $\alpha$ -thrombin (yellow) and BPTI (blue). The BPTI docking is as found in the trypsin-Arg 15-BPTI complex (Bode et al., 1984); part of the trypsin component superimposed on thrombin is given in violet color. Only part of the cleft regions of thrombin and trypsin are shown; the full structure of BPTI is given; the view is along the active-site cleft of thrombin, which points downward. Clearly visible is the collision of Trp 60D with Tyr 35I of BPTI and the overcrowding, which would be expected in the surrounding region.

namely through an exposed binding loop of canonical conformation (Huber & Bode, 1978; Laskowski & Kato, 1980; Huber & Carrell, 1989); furthermore, the interaction of this loop with thrombin seems to be restricted essentially to the immediate environment of the active-site residues (Chang et al., 1979; Hofsteenge et al., 1988). Thrombin obviously imposes particular restrictions on the shape of the serpin-binding region, which must be flat. The core of these serpin inhibitors (Löbermann et al., 1984; Baumann et al., 1991) is much larger than those of the "small" inhibitors; therefore, more intimate molecular interactions with rim elements of the thrombin active-site canyon are to be expected.

Similar geometric requirements must be met by coagulation factors susceptible to thrombin action. The thrombin interaction with fibrinogen presumably does not follow this general interaction scheme, however. Recent crystallographic studies (Martin et al., 1992; Stubbs et al., 1992) reveal that segment Phe 8–Arg 16 of the fibrinopeptide A binds in a multiturnlike conformation to the hydrophobic cleft area of the thrombin active-site region (as suggested earlier by Scheraga and coworkers). With regard to the A $\alpha$ -chain region carboxy-terminal of the scissile peptide bond Arg 16–Gly 17, one can only speculate that segment 30–44 (with several acidic residues) might bind to the fibrin(ogen) secondary binding exosite (Fig. 14) in a similar manner as seen for the bound hirudin tail (Hofsteenge et al., 1988).

Hirudin, however, interacts with thrombin in a quite different manner (Grütter et al., 1990; Rydel et al., 1990) to the canonically binding protein inhibitors (Fig. 26; see also Bode & Huber, 1991). It binds in a complementary manner with its large amino-terminal domain to thrombin-specific sites outside the enzyme's active site; only the (originally flexible [Folkers et al., 1989; Haruyama & Wuthrich, 1989]) amino-terminal segment comprising the first three hirudin residues becomes fixed under formation of a parallel  $\beta$ -pleated sheet (in contrast to PPACK [Fig. 21] and the canonical inhibitors). In addition, it associates with part of the active-site cleft and the positively charged fibrin(ogen) secondary binding exosite surface (Fig. 14) under formation of several surface-located salt

bridges and hydrophobic contacts. Thus, the thrombin-hirudin interaction gains its tremendous specificity primarily through interactions with thrombin sites not common to other serine proteinases; due to the partial occlusion of substrate-binding sites and shielding of the active-site residues (not themselves directly blocked [Rydel et al., 1991]), both macromolecular and small substrates are denied access to the thrombin active site.

## Materials and methods

### Data collection and processing

Highly purified human  $\alpha$ -thrombin was a kind gift of Drs. Jan Hofsteenge and S.R. Stone, Friedrich-Miescher-Institut, Basel. PPACK was purchased from Calbiochem or Bachem. Formation of the PPACK-human  $\alpha$ -thrombin complex and crystallization in polyethylene glycol 6000/phosphate buffer, pH 7, were described previously (Bode et al., 1989b). The crystals belong to the orthorhombic space group P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> and contain one molecule per asymmetric unit with cell constants  $a = 87.74$  Å,  $b = 67.81$  Å, and  $c = 61.07$  Å. Larger crystals initially show reflections to a minimal Bragg spacing of beyond 1.9 Å resolution. The crystals were transferred to a mother liquor of 30% polyethylene glycol 6000, buffered with 0.2 M sodium/potassium phosphate to pH 7.0 for data collection.

At first, a particularly large, but macroscopically twinned crystal obtained by microdialysis was cut into four parts; two of them were used initially for data collection on a rotation camera (see below). Only one of these crystal fragments (of size  $0.20 \times 0.4 \times 0.6$  mm<sup>3</sup>) yielded completely unsplit reflections (initially to 1.9 Å) and was transferred to a FAST television area detector diffractometer (Enraf-Nonius) using Ni-filtered CuK $\alpha$  radiation, a crystal-to-detector distance of 40 mm, and a detector angle of 10°. An X-ray data set to 2.45 Å resolution (Table 12) was collected upon rotation of the crystal fragment of 92° about its  $b$  axis and of 30° about an axis inclined by 30° to  $b$ . Capillary and crystal were kept in a cold air stream at 3 °C. The program package

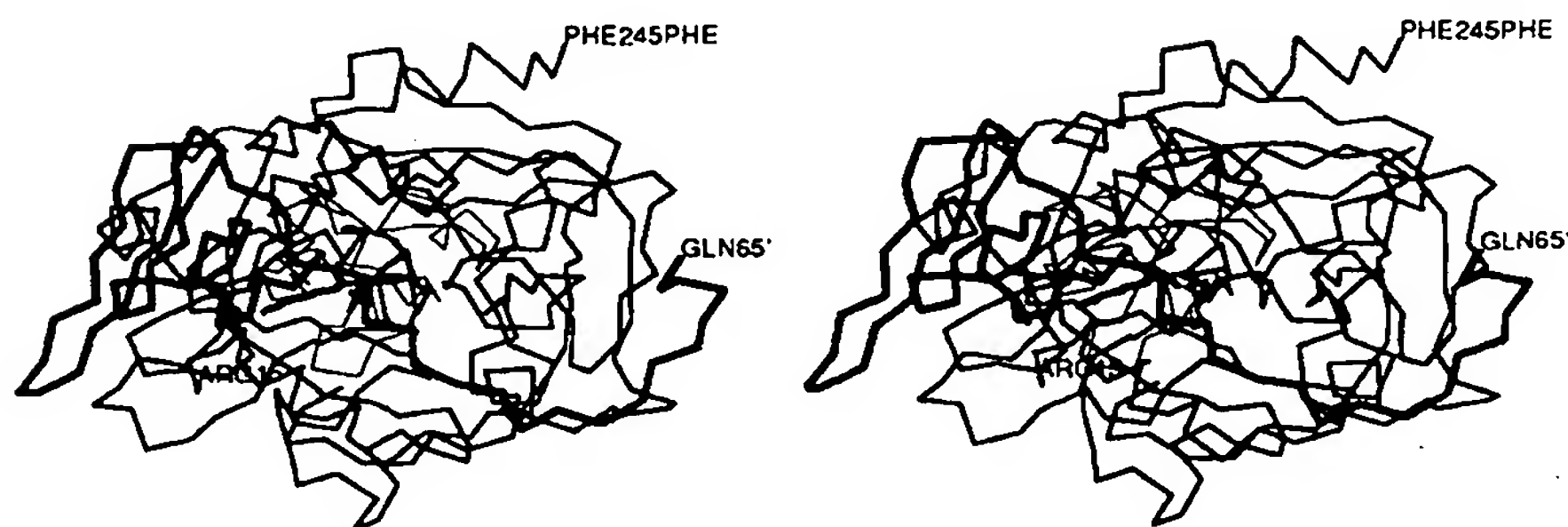


Fig. 26.  $\alpha$ -carbon structures of hirudin (bold connections) and thrombin (mediate connections) (Rydel et al., 1991), superimposed with BPTI (thin connections) as it binds to trypsin in the trypsin–Arg 15–BPTI complex (Bode et al., 1984). The P1 residue of Arg 15–BPTI, Arg 151, is shown with dashed connections.

Table 12. Intensity data statistics

A. Fast data			
Number of evaluated reflections			40,583
<i>R</i> <sub>merge</sub> <sup>a</sup>			0.109
Number of Friedel mates			18,684
<i>R</i> <sub>merge</sub> <sup>a</sup> of Friedel mates			0.047
Number of independent reflections			10,797
Maximal resolution			2.45 Å
Measured/expected reflections	316 to 2.45 Å		0.80
B. Photographic films			
Number of evaluated reflections			26,009
<i>R</i> <sub>merge</sub> <sup>a</sup>			0.071
Number of independent reflections			14,105
Maximal resolution			1.92 Å
C. Combined data			
Number of evaluated reflections			66,592
Number of independent reflections			16,910
Maximal resolution			1.92 Å
<i>R</i> <sub>merge</sub> after rejection <sup>b</sup>			0.101
Completeness			
Measured/expected reflections	316 to 2.45 Å		0.87
	2.55 to 2.45 Å		0.70
Measured/expected reflections	316 to 2.2 Å		0.77
	2.25 to 2.2 Å		0.42
Measured/expected reflections	316 to 1.92 Å		0.65
	2.05 to 1.92 Å		0.16

<sup>a</sup>  $R_{\text{merge}} = \sum h \sum i [I(i, h) - \langle I(h) \rangle] / \sum h \langle I(h) \rangle$ , where  $I(i, h)$  is the intensity of reflection  $h$  observed for the  $i$ th source, and  $\langle I(h) \rangle$  is the mean intensity of reflection  $h$  for all measurements of  $I(h)$ .

<sup>b</sup> Individual measurements were rejected if their relative deviation from the mean value  $(I - \langle I \rangle) / \langle I \rangle$  exceeded the threshold limit of 0.6.

MADNES (Messerschmidt & Pflugrath, 1987) was used to determine and refine the crystal setting parameters, crystal cell constants, and camera parameters and to evaluate the reflection data on-line. All intensities with three contiguous pixels greater than  $3.5\sigma$  above background were assigned as observed intensities. These data were corrected for absorption effects using equivalent reflections and a nonuniformity table based on previously collected data to determine the absorption ellipsoid of the crystal (Messerschmidt et al., 1990). The Friedel-mates (Table 12) were loaded, scaled, and merged using the program system PROTEIN (Steigemann, 1974). Number, quality, and completeness of these FAST data are shown in Table 12.

Additional high-resolution X-ray data up to 1.92 Å resolution (Table 12) were collected on a rotation camera (Huber, Rimsting) in the cold room at 4 °C using graphite monochromatized CuK $\alpha$  radiation. For data collection the large crystal later used for FAST data collection (see above), another slightly disordered "twin"-brother, and three additional, smaller crystals (of approximate size  $0.05 \times 0.2 \times 0.5$  mm<sup>3</sup>) were rotated about 15° each around their *b* axes. In this way the whole reciprocal space was systematically surveyed. Due to radiation damage the maximal resolution of the larger crystals de-

creased to about 2.2 Å upon X-ray exposure according to rotation photographs.

The films were scanned on an Optronics rotating-drum film scanner and evaluated with the program FILME as modified by W.S. Bennett, Jr. Reflections with intensity  $>1.0\sigma$  were further processed. Some of the film data obtained from the smaller vapor diffusion-made crystals fitted poorly to the FAST data; in spite of individual *R*<sub>sym</sub> values as low as 0.06, some of them exhibited *R*<sub>merge</sub> values (defined as  $\sum (|I - \langle I \rangle|) / \sum |I|$ ) above 0.20. Film data from one of these crystals and from all films with *R*<sub>merge</sub> values  $>0.20$  were therefore omitted from the final merged data set. Table 12 gives only the intensity statistics for accepted camera data, which represent about 70% of the asymmetric unit of reciprocal space.

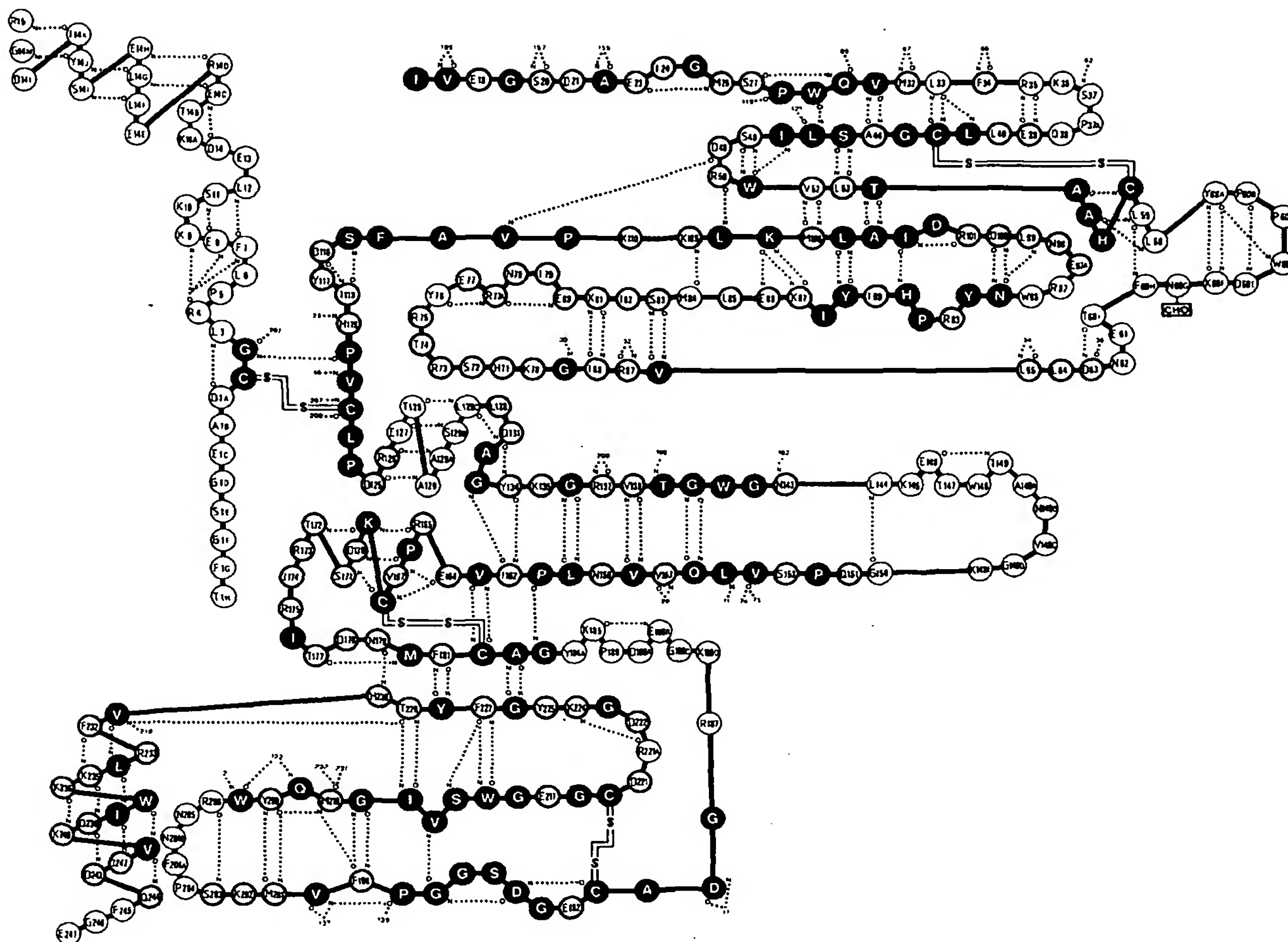
The combined and scaled data set comprises 16,910 independent X-ray data. Completeness is almost full down to 2.45 Å and gradually decreases with increasing resolution to 1.92 Å (Table 12).

#### Search for orientation and position

Figure 27 gives a schematic arrangement of the final thrombin model (see also Fig. 3). Highlighted are those amino acid residues that were selected to form the truncated model used for the search in the PPACK-thrombin crystals. Due to the anticipated closest topological similarity with thrombin, the polypeptide backbone of the bovine chymotrypsin model (Brookhaven Data Bank [Bernstein et al., 1977] codes 2CGA [Cohen et al., 1981], 4CHA [Tsukada & Blow, 1985], 5CHA [Blevins & Tulinsky, 1985]) was used as a guiding structure. Some structural features of bovine trypsin (Bode & Schwager, 1975a; Bode et al., 1976) expected to be closer at a few sites were incorporated into the truncated model as well. Ninety-seven amino acid residues of human  $\alpha$ -thrombin (emphasized by filled circles in Fig. 27) were assumed to be identical or almost identical to residues in chymotrypsin and were used in their entirety; another 105 residues (indicated by bold circles) were considered to be topologically equivalent to either chymotrypsin or trypsin and were included only with their main-chain and side-chain atoms up to C $\beta$ . No efforts were made to construct a contiguous energy-minimized search model.

In addition, the truncated model contained also the two last residues of PPACK including the hemiketal group (modeled according to the peptidyl chloromethylketone moiety of inhibited human leukocyte elastase [Wei et al., 1988]) and 10 conserved internal solvent molecules. The truncated model used for subsequent searches consisted of 1,352 active atoms, corresponding to slightly more than half of all atoms of the final peptidyl structure. This residue selection corresponds approximately to the structurally conserved regions (SCR, CR) defined by Greer (1981) and Bing et al. (1986).





**Fig. 27.** Schematic representation of the truncated model used for search; the thrombin chain segments are arranged as found for the final thrombin molecule (see Fig. 3). The backbone atoms plus  $C\beta$  of 202 residues of bovine chymotrypsin (in a few cases: of bovine trypsin, see text) assumed to be topologically equivalent to thrombin (emphasized by thick circles) were included; of these 202 residues, 97 given by filled circles were taken with full size.

The orientation of the thrombin B-chain in the thrombin crystals was determined by Patterson search methods. The truncated model was placed in a triclinic cell with orthogonal cell axes of length 120 Å. A model Patterson map was calculated with data between 8 and 3.5 Å using a B value of 20 Å<sup>2</sup>; grid points representing the 2,463 highest function values within a radial shell of 3 and 15 Å were selected. A crystal Patterson synthesis was calculated using reflections between 8 and 3.5 Å resolution on a 1-Å grid. The rotational search was carried out by calculating the product correlation function of the crystal and the modified model Patterson map as a function of the orientational angles  $\theta_1$ ,  $\theta_2$ ,  $\theta_3$  (defined as subsequent rotations around z, the new x [ $x'$ ] and the new z [ $z''$ ] axis)) in 5° steps in the angle range  $\theta_1 = 0-90^\circ$ ,  $\theta_2 = 0-180^\circ$ ,  $\theta_3 = 0-180^\circ$  (with the PROTEIN program system of

Steigemann [1974]). The rotation function showed a prominent peak of height 101.2 arbitrary units (corresponding to  $5.6\sigma$  above the mean value of 70.5) at  $10^\circ/125^\circ/90^\circ$  with second and third peak heights of 92.7 and 92.6 ( $4.1$  and  $4.1\sigma$ ), respectively. A fine scan in  $0.2^\circ$  intervals increased the value of the highest peak to 103.2 units, yielding an optimal orientation value.

The location of the thrombin B-chain with respect to the crystallographic symmetry elements was determined using the translation function and programs written by E.E. Lattmann and modified by J. Deisenhofer and R. Huber. An 8- to 3-Å model transform was calculated for the truncated search model oriented according to the rotation function solution and placed with its center of gravity at the origin of a triclinic cell with orthogonal cell axes of length 120 Å. Only one of the three cross-vector



sets yielded the highest peak on a Harker section; all three highest Harker section peaks yielded, however, a consistent solution (Table 13). The crystallographic *R*-value (defined as  $\sum (|F_{\text{obs}} - F_{\text{calc}}|) / \sum |F_{\text{obs}}|$ ) of the appropriately oriented and positioned truncated model was 0.53 for 8- to 3-Å data.

The orientation and position parameters of this model were further refined with the Fourier transform fitting program TRAREF (Huber & Schneider, 1985) resulting in rotational and positional shifts of about 3° and 0.25 Å. The *R*-value of this model was 0.493 for 8- to 3-Å data.

#### Model building and crystallographic refinement

A  $2F_{\text{obs}} - F_{\text{calc}}$  Fourier map was calculated with this model by means of PROTEIN using intensity data from

Table 13. Position and height (in arbitrary units) of highest peaks in the translation function

Vector set	Peak height	Standard deviation	Peak position			Harker section
			<i>u</i> /80	<i>v</i> /70	<i>w</i> /60	
1 → 2	382.1	58.4	40.3	55.1	30.0	<i>w</i> = 30
	374.5		8.3	2.3	0.0	
	373.9		9.6	50.3	0.0	
1 → 3	683.1	102.3	0.0	4.4	5.1	<i>u</i> = 40
	664.4		40.0	20.1	50.8	
	561.5		0.0	10.9	51.1	
1 → 4	536.6	78.8	5.7	0.0	3.5	<i>v</i> = 35
	526.6		0.0	35.0	20.3	
	430.5		5.3	0.0	59.0	

8 to 3 Å and Sim-weighted model phases. This map displayed most parts of the core-forming segments and of the PPACK moiety; in addition it showed electron density for several omitted side chains (such as Phe 11), for some insertion loops (such as loop 202–208), and for a short A-chain segment following Cys 1–Gly 2. These defined structural parts were modeled into density using the PSFRODO version (Pflugrath et al., 1984) of FRODO (Jones, 1978) on a PS390 display system (Evans & Sutherland). This new, more complete model had a decreased *R*-value of 0.44.

This model was subjected to 12 macrocycles of modeling and cycles of crystallographic refinement (Table 14). Most refinement was carried out using the energy-restrained least-squares program EREF (Jack & Levitt, 1978). Each EREF macrocycle started with inspection of Fourier and difference Fourier (and sometimes "omit") maps at the display with appropriate model manipulations and additions, followed by 5–9 refinement cycles, each consisting of calculation of structure factors and of a Sim-weighted difference Fourier (with PROTEIN), calculation of the normal equations from the difference Fourier (DERIV [Jack & Levitt, 1978]), and subsequent least-squares refinement of positional parameters with energy restraints (EREF [Jack & Levitt, 1978]). In between, thermal parameters were also refined. Most standard geometries were taken from Levitt (1974); a few improper angles were introduced to keep carboxylate and carboxamide groups planar; some peptide bond parameters were updated, and the proline parameters were changed to allow ring puckering (see Schirmer et al., 1987).

The carboxy-terminal arginine methyl residue of PPACK (Arg 31) was constructed from an EREF stan-

Table 14. Course of crystallographic refinement

Macrocycle	Active atoms	Refinement method	Resolution (Å)	Internal energy (kcal/mol)	r.m.s. bond deviation from target values	<i>R</i> -value
0	1,352	—	8 to 3	—	—	0.493
1	1,531	EREF/DERIV	8 to 3	−488	0.014	0.444
2	1,765	EREF/DERIV	8 to 2.5	−320	0.014	0.329
3	1,889	EREF/DERIV	8 to 2.5	−291	0.015	0.289
4	2,070	EREF/DERIV	6 to 2.5	−420	0.014	0.251
5	2,205	EREF/DERIV	6 to 2.5	−472	0.016	0.234
6	2,331	EREF/DERIV	6 to 2.2	−752	0.016	0.234
7	2,365	EREF/DERIV	6 to 1.92	−1,056	0.018	0.200
8	2,391	EREF/DERIV	6 to 1.92	−1,362	0.018	0.178
9	2,589	EREF/DERIV	6 to 1.92	−1,298	0.018	0.173
10	2,650	EREF/DERIV	6 to 1.92	−1,227	0.019	0.171
11	2,833	X-PLOR	8 to 1.92	−2,857	0.013	0.174
12a	2,819	X-PLOR	8 to 1.92	−3,180	0.013	0.156
12b	2,833	EREF	8 to 1.92	−1,442	0.018	0.159

dard arginine residue by adding a methylene group and enforcing a tetrahedral geometry at the carbonyl carbon as done for the complex of human leukocyte elastase with a peptidyl chloromethylketone (Wei et al., 1988). All non-bonded restraint parameters for Ser 195 O $\gamma$  and for the Arg 31 methylene group were set to zero to allow their unbiased approach to the Arg 31 carbonyl group and to the His 57 N $\epsilon$ 2 atom, respectively. At later stages of refinement, solvent molecules interpreted as water molecules were inserted at stereochemically reasonable positions if the difference electron density exceeded  $4\sigma$  (using PROTEIN routines).

The graphic program MAIN (D. Turk) was employed in macrocycle 11 in combination with the conjugate gradient minimization procedure of X-PLOR version 1.5 (Brünger et al., 1987). Crystallographic refinement with X-PLOR involved 50 cycles of positional refinement followed by 50 steps of B-value optimization. New solvent molecules were searched for by means of MAIN routines that checked the difference Fourier map for peaks greater than  $2.5\sigma$  and inserted waters (closest to the original molecule) if located within hydrogen-bond distance (1.8–3.5 Å) of polar protein atoms or of other solvent molecules; current solvent molecules were checked for height of density in the Fourier map and were deleted if positioned in density less than  $0.8\sigma$  above the mean value. The X-PLOR force-field parameters given in PARAM19X.PRO (which treats only polar hydrogens explicitly) were used for restrained crystallographic refinement; all side-chain charges were set to zero (as recommended in the X-PLOR manual), and all histidine residues were treated as singly protonated, with His 57 carrying a proton at N $\delta$ 1 and all other histidines at N $\epsilon$ 2.

PPACK residue Arg 31 was modeled according to an X-PLOR standard arginine residue; additional patch residues were built to allow covalent linkages to the preceding residue (Pro 21), via the carbonyl carbon (CF) to Ser 195 O $\gamma$  and via its methylene carbon (CJ) to His 57 N $\epsilon$ 2, respectively. The hemiketal group of Arg 31 was restrained to an ideal tetrahedral state (with all angles around the carbonyl carbon [CF] being  $109.5^\circ$ ); the lengths of both CF-O bonds (from the carbonyl carbon to its oxygen and to Ser 195 O $\gamma$ ) were restrained to 1.42 Å (as C-O single bonds), and the bond length between the methylene carbon and His 57 N $\epsilon$ 2 to 1.49 Å (taken from X-PLOR parameters).

The course of the crystallographic refinement is shown in Table 14. The density calculated with the first, already improved model displayed spurious density for most of the insertion loops, which only from macrocycle 3 on became clear enough to be modeled. From macrocycle 4 on, the full B-chain was present; several peptide segments, in particular some insertion loops and the carboxy-terminus, remained subject to changes until the very end of refinement. From macrocycle 6 on, additional water molecules

were localized and included in refinement. At macrocycle 8 the A-chain was clearly traceable from Asp 1A to Tyr 14J; the preceding residues Ala 1B and Glu 1C were modeled and refined with relatively bad conformations; handicapped by several breaks in density, the first five residues, Thr 1H to Gly 1D, were tentatively placed along a shallow groove along the B-chain surface. At this stage no continuous density was visible for the last four residues, Ile 14K–Arg 15, of the A-chain.

The model resulting from macrocycle 10 was that described recently (Bode et al., 1989b). It comprised 290 defined residues (lacking the four carboxy-terminal residues of the A-chain and the carboxy-terminal Glu 247 of the B-chain) with some side chains remaining partially undefined, and 291 solvent molecules, and had an *R*-value of 0.171 (6–1.92-Å data).

Upon further inspection of the maps it became clear that some hitherto undefined side chains could be localized and furthermore that Lys 110 (with a very mobile side chain) had to be rotated and that the density “tongue” extending away from the main chain and being formerly interpreted as the Lys 110 side chain should rather accommodate its carbonyl group; as a consequence, the peptide bond formed with the succeeding prolyl residue, Pro 111, should (contrary to our previous suggestion) be in a trans rather than in a cis conformation. Thus, only one of the proline residues of human thrombin (Pro 37) is in a cis conformation, in full agreement with the structures of the thrombin–hirudin complex (Rydel et al., 1991). In addition, a slightly different pathway for the A-chain segment preceding Glu 1C was found; as a result, the peptide segment from Thr 1H to Glu 1C is still arranged in a multiple-turn conformation along the same B-chain surface as before, but is now slightly better defined by density with a more reasonable stereochemistry. Likewise, some faint but (at very low contouring) continuous density allowed the placement of the carboxy-terminal four A-chain residues and Glu 247 of the B-chain with more confidence and to include all side-chain atoms assigned before as dummy atoms in phase calculation.

The *R*-value of this partially revised thrombin model was 0.193. A subsequent restrained refinement with X-PLOR (macrocycle 11 [Table 14]) using the default bond and angle B-value restraints brought it down to 0.174. In a final unrestrained B-value refinement with X-PLOR (macrocycle 12a) or EREF (macrocycle 12b) this model converged (without rejection of any bad reflections) to final *R*-values of 0.156 and 0.159, respectively. The final parameters of both models obtained from X-PLOR and EREF refinement are given in Table 15. The models are virtually identical (with an r.m.s. deviation of 0.07 Å between all 295  $\alpha$ -carbon atoms, i.e., well below the estimated accuracy of each model); the small differences result from different target and restraint parameters in both procedures.

Table 15. Final model parameters of PPACK-thrombin

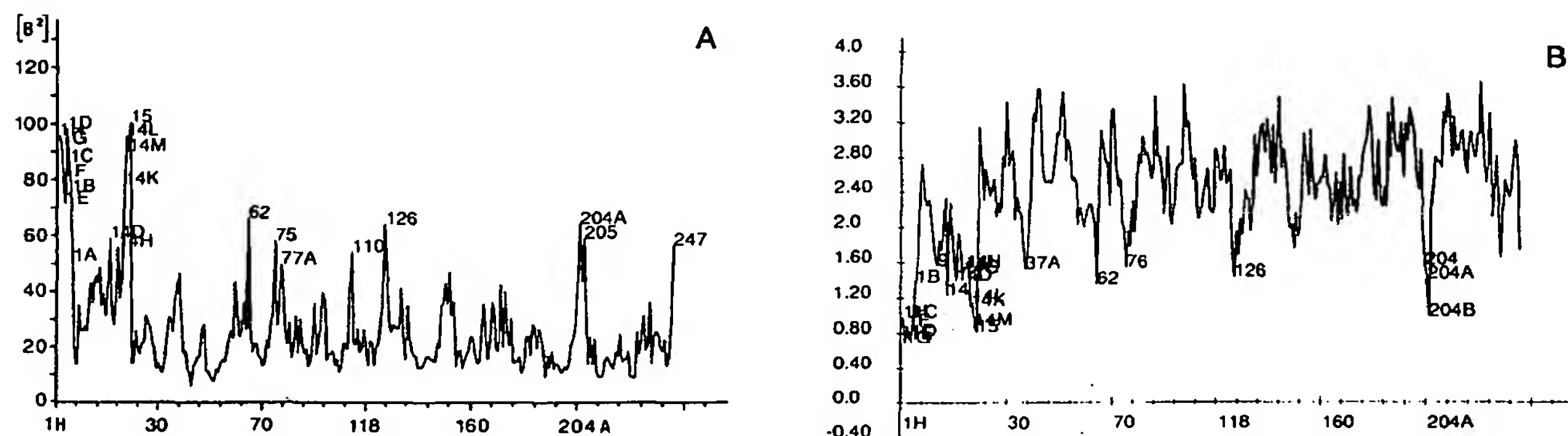
	EREF		X-PLOR
Number of active thrombin atoms		2,380	
Number of active PPACK atoms		30	
Number of active solvent atoms	423		409
r.m.s. standard deviation from target values			
Bond lengths	0.018 Å		0.013 Å
Bond angles	3.085°		3.137°
Total energy	-1,442 kcal/mol		-3,733 kcal/mol
Resolution range		8 to 1.92 Å	
Number of reflections used in refinement		15,019	
R value	0.159		0.156
Mean B value			
Whole data set	26.8 Å <sup>2</sup>		29.9 Å <sup>2</sup>
Thrombin nonhydrogen atoms	22.9 Å <sup>2</sup>		26.8 Å <sup>2</sup>
PPACK nonhydrogen atoms	8.4 Å <sup>2</sup>		10.2 Å <sup>2</sup>
Solvent molecules	45.3 Å <sup>2</sup>		49.7 Å <sup>2</sup>
Estimated mean coordinate error from agreement between observed and calculated structure factor amplitudes			
According to Luzzati (1952)	0.23 Å		0.22 Å
According to Cruickshank (1949)	0.11 Å		0.11 Å

The B-values for each thrombin residue, averaged over all nonhydrogen atoms, are shown in Figure 28A; the mean electron density per residue is given in Figure 28B; with the exception of peptide bond 14L-14M ( $<0.6\sigma$ ) and 1H-1G and 1F-1E ( $<0.8\sigma$ ) the main chain is everywhere represented by continuous electron density above  $1\sigma$ . In general, the oscillations reflect the alternate arrangement on the surface and in the interior of the molecule; particularly high B-values are observed for both A-chain termini, for the exposed B-chain loops 70-80 and 204A-205, and for the carboxy-terminus of the B-chain.

A very few noninterpreted maxima and minima re-

main in the solvent region of the final difference Fourier map with heights up to  $5\sigma$  and  $-3\sigma$ , respectively; no new solvent molecules could be placed at stereochemically reasonable positions, however. According to the Ramachandran plot (Fig. 29) almost all main-chain angles of the thrombin model are within the allowed regions defined by Ramakrishnan and Ramachandran (1965). The angle pairs outside belong to the weakly defined amino- and carboxy-termini of the thrombin A-chain; they are labeled in Figure 29 and presumably reflect multiple conformations of the respective residues.

For such energy-restrained least-squares procedures it



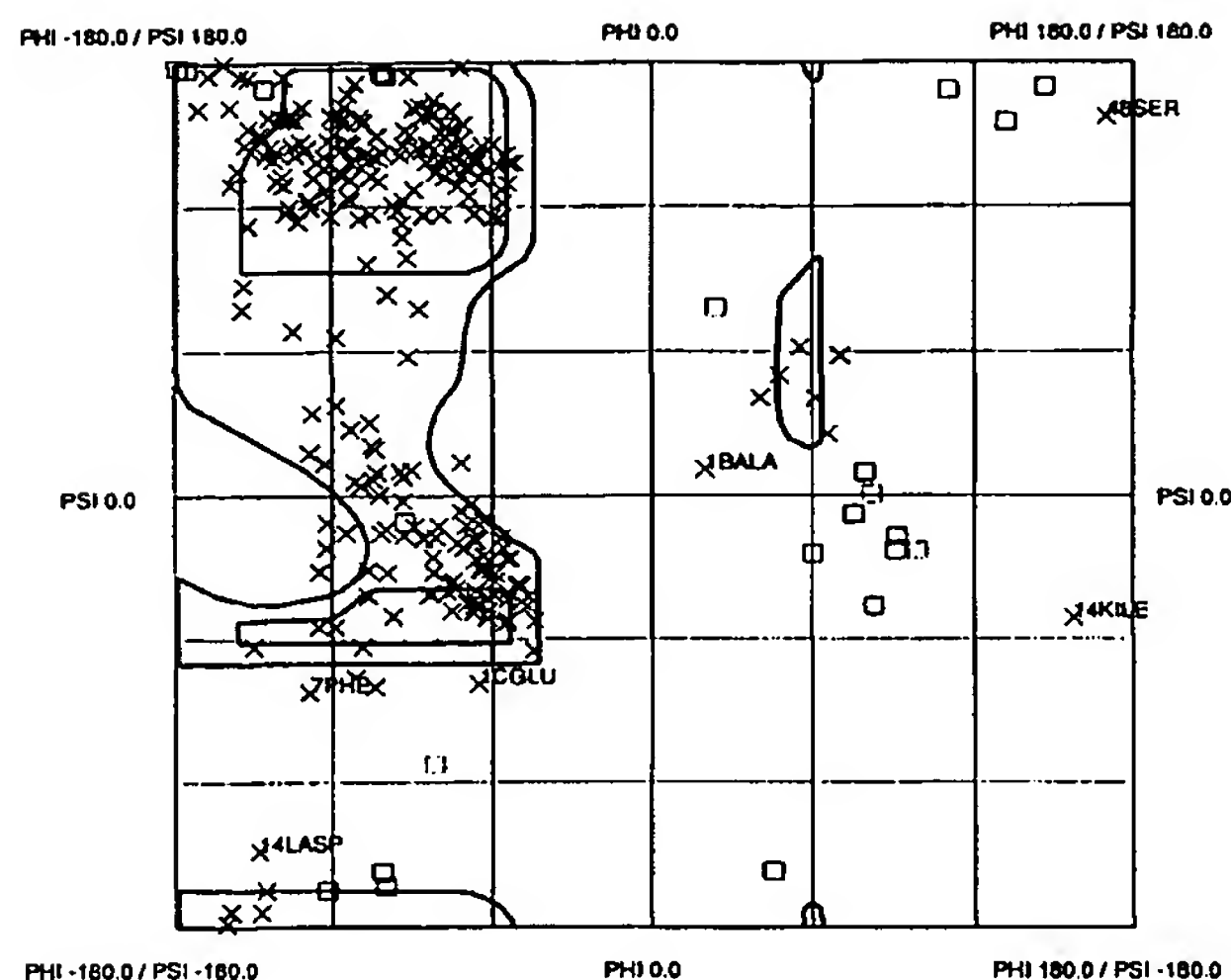


Fig. 29. Ramachandran plot of the refined thrombin model. O, glycyl residues; X, all other residues.

is difficult to define an adequate standard deviation for positional parameters of the atoms. From the scattering angle dependence of the  $R$ -values (Luzzati, 1952) the upper limit of the mean positional error of the refined model is estimated to about 0.21 Å (Fig. 30; Table 15). The formula derived by Cruickshank (1949) gives a lower limit of 0.11 Å for the accuracy of the final structure.

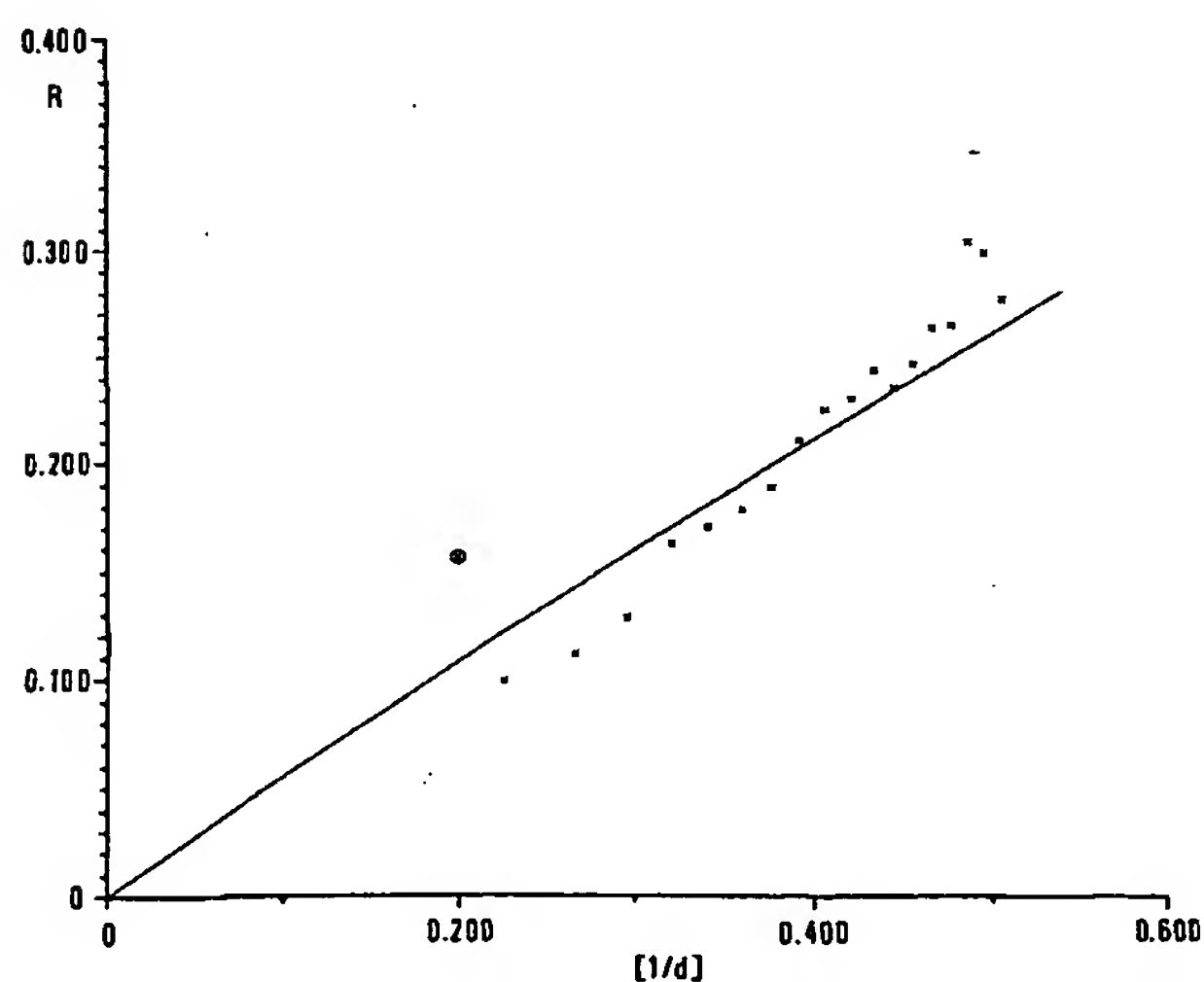


Fig. 30. Luzzati plot of the final thrombin model after X-PLOR refinement.

### Hydrogen-bond definition, alignment procedures, and electrostatic calculations

Accessible surface calculations and identification of possible hydrogen bonds were carried out by means of the program HYBAC (Levitt, 1974). Possible hydrogen-bond donor-acceptor interactions were considered as hydrogen bonds if the interaction energy was  $E < -0.7$  kcal/mol calculated with DSSP according to Kabsch and Sander (1983). Interface areas were calculated with a program of T.J. Richmond (see Richards, 1977). Most of the plots were made using the graphics facilities of FRODO, MAIN (D. Turk), and a RIBBON version (Priestle, 1988) modified by one of us (A.K.)

Optimal superposition of related proteinases and determination of the number of topologically equivalent  $\alpha$ -carbon atoms was achieved with the program OVLAP (Rossmann & Argos, 1975) as modified by W.S. Bennett; the following selection parameters were used:  $P_{\text{cut}} = 0.1$ ;  $E_1 = E_2 = 2.0$ ; run length  $> 6$ .

The amino acid sequences were aligned using the program system ALIGN (Needleman & Wunsch, 1970). In single runs only pairs of sequences were compared; the Mutation Data Matrix (250 PAMS) was used as scoring matrix; a penalty of 20 was assigned to each break, and 100 random runs were performed for each comparison. Protein sequence searches were done against the MIPSX merged Protein Sequence Data Base, Release 25.0, June 30, 1990 (Martinsried Institute for Protein Sequences).

The electrostatic calculations were performed on the basis of the modified Tanford-Kirkwood theory (Tanford & Kirkwood, 1957; Karshikov et al., 1989) as well as of the finite-difference algorithm developed by Gilson et al. (1985) and Klapper et al. (1986). In both methods the charges carried by the titratable groups and the peptide dipoles are considered. The point charges were attributed (1) to those between NH1 and 2 of the arginine guanidyl groups, (2) to N $\epsilon$  of lysine residues, (3) to the center of imidazole rings, and (4) to those between both carboxylate oxygen atoms of aspartates and glutamates; orientation and size of the peptide dipoles were taken as described by Karshikov et al. (1989). The solvent accessibilities of the charged groups in the protein structures defined as the accessibility ratio to that of a fully extended Ala-Xaa-Ala peptide were calculated according to Lee and Richards (1981). Dielectric constants of 4 and 78 were used for the space inside and outside of the protein. All calculations were performed for an ionic strength of 0.125 and a pH value of 7.0. In the case of the Tanford-Kirkwood model the intrinsic pK values were taken from Matthew (1985).

Both methods yielded quite consistent results regarding the shape of the electric potential; the contouring surfaces shown in Figure 8 and the electrostatic interaction energies presented in this paper are those derived by



means of the Tanford-Kirkwood model. In the case of the finite-difference method, the calculations were started with coulombic boundary conditions followed by two steps of focusing calculations. The electrostatic equipotential surfaces were programmed (A.K.) to be displayed on a PS system using the FRODO graphics facilities.

### Acknowledgments

We thank Prof. Dr. Robert Huber in particular for his long-standing interest in this work, Drs. J. Hofsteenge and S.R. Stone, Friedrich-Miescher-Institut, for human thrombin samples and for stimulating discussions, Dr. M. Stubbs for careful reviewing of this manuscript, and Drs. P. Ruecknagel, E. Jäger, and H. Oschkinat for preparation of and for NMR measurements with the YPPWDN peptide. The help of I. Mayr, U. Baumann, and other colleagues of the Abteilung für Struktur-forschung and of the Martinsried computer center is acknowledged. The financial support by the Sonderforschungsbereich 207 (H1) of the University of Munich, by the Fonds der Chemischen Industrie, by the Alexander-von-Humboldt-Stiftung, and by the Research Council of Slovenia/Ministry for Science and Technology is gratefully acknowledged.

### References

- Ascenzi, P., Coletta, M., Amiconi, G., de Cristofaro, R., Bolognesi, M., Guarneri, M., & Menegatti, E. (1988). Binding of the bovine basic pancreatic trypsin inhibitor (Kunitz) to human  $\alpha$ -,  $\beta$ - and  $\gamma$ -thrombin; a kinetic and thermodynamic study. *Biochim. Biophys. Acta* 956, 156-161.
- Bajusz, S., Barabas, E., Tolnay, P., Szell, E., & Bagdy, D. (1978). Inhibition of thrombin and trypsin by tripeptide aldehydes. *Int. J. Peptide Protein Res.* 123, 217-221.
- Bajusz, S., Szell, E., Bagdy, D., Barabas, E., Horvath, G., Dioszegi, M., Fittler, Z., Szabo, G., Juhasz, A., Tomori, E., & Szilagy, E. (1990). Highly active and selective anticoagulants: D-Phe-Pro-Arg-H, a free tripeptide aldehyde prone to spontaneous inactivation, and its stable N-methyl derivative, D-Me-Phe-Pro-Arg-H. *J. Mol. Chem.* 33, 1729-1735.
- Bar-Shavit, R., Eldor, A., & Vlodavsky, I. (1989). Binding of thrombin to subendothelial extracellular matrix; protection and expression of functional properties. *J. Clin. Invest.* 84, 1096-1104.
- Bar-Shavit, R., Kahn, A., Mann, K.G., & Wilner, G.D. (1986). Growth-promoting effects of esterolytically inactive thrombin on macrophages. *Cell Biochem.* 32, 261-272.
- Bar-Shavit, R., Kahn, A., Mudd, M.S., Wilner, G.D., Mann, K.G., & Fenton, J.W., II (1984). Localization of a chemotactic domain in human thrombin. *Biochemistry* 23, 397-400.
- Bar-Shavit, R., Kahn, A., Wilner, G.D., & Fenton, J.W., II (1983). Monocyte chemotaxis: Stimulation by a specific exosite region in thrombin. *Science* 220, 728-731.
- Bar-Shavit, R., Sabbah, V., Lampugnani, M.G., Marchisio, P.C., Fenton, J.W., II, Vlodavsky, I., & Dejana, E. (1991). An Arg-Gly-Asp sequence within thrombin promotes endothelial cell adhesion. *J. Cell Biol.* 112, 335-344.
- Baumann, U., Huber, R., Bode, W., Grosse, D., Lesjak, M., & Laurell, C.B. (1991). Crystal structure of cleaved human  $\alpha_1$ -antichymotrypsin at 2.7 Å resolution and its comparison with other serpins. *J. Mol. Biol.* 218, 595-606.
- Berg, W., Hillvärn, B., Arwin, H., Stenberg, M., & Lundström, I. (1979). The isoelectric point of thrombin and its behaviour compared to prothrombin at some solid surfaces. *Thromb. Haemostasis* 42, 972-982.
- Berliner, L.J., Birktoft, J.J., Miller, T.L., Musci, G., Scheffler, J.W., Shen, Y.Y., & Sugawara, Y. (1986). Thrombin: Active-site topography. *Ann. N.Y. Acad. Sci.* 485, 80-95.
- Berliner, L.J. & Shen, Y.Y.L. (1977). Physical evidence for an apolar binding site near the catalytic center of human  $\alpha$ -thrombin. *Biochemistry* 16, 4622-4626.
- Berliner, L.J., Sugawara, Y., & Fenton, J.W., II (1985). Human  $\alpha$ -thrombin binding to nonpolymerized fibrin-sepharose: Evidence for an anionic binding region. *Biochemistry* 24, 7005-7009.
- Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rogers, J.R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977). The protein data bank: A computer-based archival file for macromolecular structures. *J. Mol. Biol.* 122, 535-542.
- Bezaud, A. & Guillin, M.C. (1988). Enzymic and nonenzymic properties of human  $\beta$ -thrombin. *J. Biol. Chem.* 263, 3576-3581.
- Bing, D.H., Cory, M., & Fenton, J.W., II (1977). Exosite affinity labeling of human thrombins; similar labeling on the A chain and B chain fragments of clotting  $\alpha$ - and non-clotting  $\gamma/\beta$  thrombins. *J. Biol. Chem.* 252, 8027-8034.
- Bing, D.H., Feldmann, R.J., & Fenton, J.W., II (1986). Structure and function relationships of thrombin based on the computer generated three-dimensional model of the B chain of bovine thrombin. *Ann. N.Y. Acad. Sci.* 485, 104-119.
- Bizios, R., Lai, L., Fenton, J.W., II, & Malik, A.B. (1986). Thrombin-induced chemotaxis and aggregation of neutrophils. *J. Cell. Physiol.* 128, 485-490.
- Björk, I. & Lindahl, V. (1982). Mechanism of the anticoagulant action of heparin. *Mol. Cell. Biochem.* 48, 161-182.
- Blevins, R.A. & Tulinsky, A. (1985). The refinement and the structure of the dimer of  $\alpha$ -chymotrypsin at 1.67 Å resolution. *J. Biol. Chem.* 260, 4264-4275.
- Blombäck, B., Blombäck, B., Hessel, B., & Iwanaga, S. (1967). Structure of N-terminal fragments of fibrinogen and specificity of thrombin. *Nature* 215, 1445-1448.
- Blombäck, B., Hessel, H., Hogg, D., & Claesson, G. (1977). Substrate specificity of thrombin on proteins and synthetic substrates. In *Chemistry and Biology of Thrombin* (Lundblad, R.L., Fenton, J.W., II, & Mann, K.G., Eds.), pp. 275-290. Ann Arbor Science Publishers, Ann Arbor, Michigan.
- Bode, W. (1979). The transition of bovine trypsinogen to a trypsin-like state upon strong ligand binding. II. The binding of the pancreatic trypsin inhibitor and of isoleucine-valine and of sequentially related peptides to trypsinogen and to p-guanidinobenzoate-trypsinogen. *J. Mol. Biol.* 127, 357-374.
- Bode, W., Chen, Z., Bartels, K., Kutzbach, C., Schmidt-Kastner, G., & Bartunik, H. (1983). Refined 2 Å X-ray crystal structure of porcine pancreatic kallikrein A, a specific trypsin-like serine proteinase. Crystallization, structure determination, crystallographic refinement, structure and its comparison with bovine trypsin. *J. Mol. Biol.* 164, 237-282.
- Bode, W., Greyling, H.J., Huber, R., Otlewski, J., & Wilusz, T. (1989a). The refined 2.0 Å X-ray crystal structure of the complex formed between bovine  $\beta$ -trypsin and CMTI-I, a trypsin inhibitor from squash seeds (*Cucurbita maxima*). Topological similarity of the squash inhibitors with the carboxypeptidase A inhibitor from potatoes. *FEBS Lett.* 242, 285-292.
- Bode, W. & Huber, R. (1986). Crystal structures of pancreatic serine endopeptidases. In *Molecular and Cellular Basis of Digestion* (Desnuelle, P., Ed.), pp. 213-234. Elsevier, Amsterdam, New York, Oxford.
- Bode, W. & Huber, R. (1991). Ligand binding: Proteinase-protein inhibitor interactions. *Curr. Opin. Struct. Biol.* 1, 45-52.
- Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S.R., & Hofsteenge, J. (1989b). The refined 1.9 Å crystal structure of human  $\alpha$ -thrombin: Interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment. *EMBO J.* 8, 3467-3475.
- Bode, W., Papamokos, E., & Musil, D. (1987). The high resolution X-ray crystal structure of the complex formed between subtilisin Carlsberg and the inhibitor eglin C, an elastase inhibitor from the leech *Hirudo medicinalis*. Structural analysis, subtilisin structure and interface geometry. *Eur. J. Biochem.* 166, 673-692.
- Bode, W. & Schwager, P. (1975a). The refined crystal structure of bovine  $\beta$ -trypsin at 1.8 Å resolution. II. Crystallographic refinement,

- calcium binding site, benzamidine binding site and active site at pH 7. *J. Mol. Biol.* 98, 693-717.
- Bode, W. & Schwager, P. (1975b). The single calcium binding site of crystalline bovine  $\beta$ -trypsin. *FEBS Lett.* 56, 139-143.
- Bode, W., Schwager, P., & Huber, R. (1976). Structural studies on the pancreatic trypsin inhibitor-trypsin complex and its free components: Structure and function relationship in serine protease inhibition and catalysis. In *Proteolysis and Physiological Regulation*, pp. 43-76. Miami Winter Symposia. Academic Press, New York, San Francisco, London.
- Bode, W., Schwager, P., & Huber, R. (1978). The transition of bovine trypsinogen to a trypsin-like state upon strong ligand binding. The refined crystal structures of bovine trypsinogen-pancreatic inhibitor complex and of its ternary complex with Ile-Val at 1.9 Å resolution. *J. Mol. Biol.* 118, 99-112.
- Bode, W., Turk, D., & Stürzebecher, J. (1990). Geometry of binding of the benzamidine- and arginine-based inhibitors NAPAP and MQPA to human  $\alpha$ -thrombin. X-ray crystallographic determination of the NAPAP-trypsin complex and modeling of NAPAP-thrombin and MQPA-thrombin. *Eur. J. Biochem.* 193, 175-182.
- Bode, W., Walter, J., Huber, R., Wenzel, H.R., & Tschesche, H. (1984). The refined 2.2 Å (0.22 nm) X-ray crystal structure of the ternary complex formed by bovine trypsinogen, valine-valine, and the Arg 15 analogue of bovine pancreatic trypsin inhibitor. *Eur. J. Biochem.* 144, 185-190.
- Bode, W., Wei, A.Z., Huber, R., Meyer, E., Travis, J., & Neumann, S. (1986). X-ray crystal structure of the complex of human leukocyte elastase (PMN elastase) and the third domain of the turkey ovomucoid inhibitor. *EMBO J.* 5, 2453-2458.
- Boissel, J.P., LeBonniec, B., Rabet, M.J., Labie, D., & Elion, J. (1984). Covalent structures of  $\beta$  and  $\gamma$  autolytic derivatives of human  $\alpha$ -thrombin. *J. Biol. Chem.* 259, 5691-5697.
- Brandstetter, H., Turk, D., Hoeffken, W., Grosse, D., Stürzebecher, J., Martin, P.D., Edwards, B.F.P., & Bode, W. (1992). X-ray crystal structure of thrombin complexes with the benzamidine- and arginine-based inhibitors NAPAP, 4-TAPAP and MQPA: A starting point for elaborating improved antithrombotics. *J. Mol. Biol.* (in press).
- Braun, P.J., Hofsteenge, J., Chang, J.Y., & Stone, S.R. (1988). Preparation and characterization of proteolyzed forms of human  $\alpha$ -thrombin. *Thromb. Res.* 50, 273-283.
- Brezniak, D.V., Brower, M.S., Witting, J.I., Walz, D.A., & Fenton, J.W., II (1990). Human  $\alpha$ - to  $\beta$ -thrombin cleavage occurs with neutrophil cathepsin G or chymotrypsin while fibrinogen clotting activity is retained. *Biochemistry* 29, 3526-3542.
- Brower, M.S., Walz, D.A., Garry, E.E., & Fenton, J.W., II (1987). Thrombin-induced elastase alters human  $\alpha$ -thrombin function: Limited proteolysis near the  $\gamma$ -cleavage site results in decreased fibrinogen clotting and platelet stimulatory activity. *Blood* 69, 813-819.
- Brünger, A.T., Kuriyan, K., & Karplus, M. (1987). Crystallographic R factor refinement by molecular dynamics. *Science* 235, 458-460.
- Burley, S.K. & Petsko, G.A. (1985). Aromatic-aromatic interaction: A mechanism of protein structure stabilization. *Science* 229, 23-28.
- Butkowski, R.J., Elion, J., Downing, M.R., & Mann, K.G. (1977). Primary structure of human prothrombin 2 and  $\alpha$ -thrombin. *J. Biol. Chem.* 252, 4942-4957.
- Cardin, A.D. & Weintraub, H.J.R. (1989). Molecular modeling of protein-glucosaminoglycan interactions. *Arteriosclerosis* 9, 21-32.
- Carney, D.E., Herbosa, G.J., Stiernberg, J., Bergmann, J.S., Gordon, E.A., Scott, D.L., & Fenton, J.W., II (1986). Double signal hypothesis for thrombin initiation of cell proliferation. *Semin. Thromb. Hemostasis* 12, 231-240.
- Chang, J.Y. (1986). The structures and proteolytic specificities of autolysed human thrombin. *Biochem. J.* 240, 797-802.
- Chang, J.-Y. (1989). The hirudin-binding site of human  $\alpha$ -thrombin; identification of lysyl residues which participate in the combining site of hirudin-thrombin complex. *J. Biol. Chem.* 264, 7141-7146.
- Chang, T.L., Feinman, R.D., Laudis, B.H., & Fenton, J.W., II (1979). Antithrombin reactions with  $\alpha$ - and  $\gamma$ -thrombins. *Biochemistry* 18, 113-119.
- Charo, I.F., Bekeart, L.S., & Phillips, D.R. (1987). Platelet glycoprotein IIb-IIIa-like proteins mediate endothelial cell attachment to adhesive proteins and the extracellular matrix. *J. Biol. Chem.* 262, 9935-9938.
- Chow, M.M., Meyer, E.F., Jr., Bode, W., Kam, C.-M., Radhakrishnan, R., Vijayalakshmi, J., & Powers, J.C. (1990). The 2.2 Å resolution X-ray crystal structure of the complex of trypsin inhibited by 4-chloro-3-ethoxy-7-guanidinocoumarin: A proposed model of the thrombin-inhibitor complex. *J. Am. Chem. Soc.* 112, 7783-7789.
- Church, F.C., Pratt, C.W., Noyes, C.M., Kalayanamit, T., Sherril, G.B., Tobin, R.B., & Meade, B. (1989). Structural and functional properties of human  $\alpha$ -thrombin, phosphopyridoxylated  $\alpha$ -thrombin and  $\gamma$ -thrombin. Identification of lysyl residues in  $\alpha$ -thrombin that are critical for heparin and fibrin(ogen) interactions. *J. Biol. Chem.* 264, 18419-18425.
- Cohen, H.G., Silverton, E.W., & Davis, D. (1981). Refined crystal structure of  $\gamma$ -chymotrypsin at 1.9 Å resolution. Comparison with other pancreatic serine proteases. *J. Mol. Biol.* 148, 449-479.
- Cruickshank, D.W.J. (1949). The accuracy of electron-density maps in X-ray analysis with special reference to dibenzyl. *Acta Crystallogr.* 2, 65-82.
- Cunningham, D.D. & Farrell, D.H. (1986). Thrombin interactions with cultured fibroblasts: Relationship to mitogenic stimulation. *Ann. N.Y. Acad. Sci.* 485, 240-248.
- De Cristofaro, R. & DiCera, E. (1990). Effect of protons on the amidase activity of human  $\alpha$ -thrombin. *J. Mol. Biol.* 216, 1077-1085.
- Degen, S.J.F., MacGillivray, R.T.A., & Davie, E.W. (1983). Characterization of the complementary deoxyribonucleic acid and gene coding for human prothrombin. *Biochemistry* 22, 2087-2097.
- Dihanich, M. & Monard, M. (1990). cDNA sequence of rat prothrombin. *Nucleic Acids Res.* 18, 4251.
- Dotz, J., Köhler, S., & Baici, A. (1988). Interaction of site specific hirudin variants with  $\alpha$ -thrombin. *FEBS Lett.* 229, 87-90.
- Doyle, M.F. & Mann, K.G. (1990). Multiple active forms of thrombin IV. Relative activities of meizothrombins. *J. Biol. Chem.* 265, 10693-10710.
- Elion, J., Boissel, J.P., LeBonniec, B., Bezeaud, A., Jandrot-Perrus, M., Rabet, M.J., & Guillin, M.C. (1986). Proteolytic derivatives of thrombin. *Ann. N.Y. Acad. Sci.* 485, 16-26.
- Esmon, N.L., Carrol, R.C., & Esmon, C.T. (1983). Thrombomodulin blocks the ability of thrombin to activate platelets. *J. Biol. Chem.* 258, 12238-12242.
- Esmon, N.L., Owen, W.G., & Esmon, C.T. (1982). Isolation of membrane-bound cofactor for thrombin-catalyzed activation of protein C. *J. Biol. Chem.* 257, 859-864.
- Fenton, J.W., II (1981). Thrombin specificity. *Ann. N.Y. Acad. Sci.* 370, 468-495.
- Fenton, J.W., II (1986). Thrombin. *Ann. N.Y. Acad. Sci.* 485, 5-15.
- Fenton, J.W., II (1988). Regulation of thrombin generation and functions. *Semin. Thromb. Hemostasis* 14, 234-240.
- Fenton, J.W., II, Fasco, M.J., Stackrow, A.B., Aronson, D.L., Young, A.M., & Finlayson, J.S. (1977). Human thrombins: Production, evaluation and properties of  $\alpha$ -thrombin. *J. Biol. Chem.* 252, 3587-3598.
- Fenton, J.W., II, Olson, T.A., Zabinski, M.P., & Wilner, G.D. (1988). Anion binding exosite of human  $\alpha$ -thrombin and fibrin(ogen) recognition. *Biochemistry* 27, 7106-7112.
- Fenton, J.W., II, Witting, J.I., Pouliott, C., & Fareed, J. (1989). Thrombin anion-binding exosite interactions with heparin and various polyanions. *Ann. N.Y. Acad. Sci.* 556, 158-165.
- Fenton, J.W., II, Zabinski, M.P., Hsieh, K., & Wilner, G.D. (1981). Thrombin non-covalent protein binding and fibrin(ogen) recognition. *Thromb. Haemostasis* 46, 177.
- Folkers, P.J.M., Clore, G.M., Driscoll, P.C., Dotz, J., Köhler, S., & Gronenborn, A.M. (1989). Solution structure of recombinant hirudin and the Lys-47-Glu mutant: A nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing study. *Biochemistry* 28, 2601-2617.
- Freer, S.T., Kraut, J., Robertus, J.D., Wright, H.T., & Xuong, N.H. (1970). Chymotrypsinogen: 2.5 Å Crystal structure, comparison with  $\alpha$ -chymotrypsin, and implications for zymogen activation. *Biochemistry* 9, 1997-2009.
- Frieze-Degen, S.J., Schaefer, L.A., Jamison, C.S., Grant, S.G., Fitzgibbon, J.J., Pai, J.A., Chapman, V.M., & Elliott, R.W. (1990). Characterization of the cDNA coding for mouse prothrombin and localization of the gene on mouse chromosome 2. *DNA Cell Biol.* 9, 487-498.
- Gilson, M.K., Rashin, A., Fine, R., & Honig, B. (1985). On the calculation of electrostatic interactions in proteins. *J. Mol. Biol.* 183, 503-516.



- Glenn, K.C., Frost, G.H., Bergmann, J.S., & Carney, D.H. (1988). Synthetic peptides bind to high-affinity thrombin receptors and modulate thrombin mitogenesis. *Peptide Res.* 1, 65-73.
- Greer, J. (1981). Comparative model-building of the mammalian serine proteinases. *J. Mol. Biol.* 153, 1027-1042.
- Griffith, M.J. (1982). Kinetics of the heparin-enhanced antithrombin III/thrombin reaction. Evidence for a template model for the mechanism of action of heparin. *J. Biol. Chem.* 257, 7360-7365.
- Griffith, M.J., Kingdon, H.S., & Lundblad, R.L. (1979). The interaction of heparin with human  $\alpha$ -thrombin: Effect on the hydrolysis of anilide tripeptide substrates. *Arch. Biochem. Biophys.* 195, 378-384.
- Grütter, M.G., Priestle, J.P., Rahuel, J., Grossenbacher, H., Bode, W., Hofsteenge, J., & Stone, S.R. (1990). Crystal structure of the thrombin-hirudin complex: A novel mode of serine protease inhibitor. *EMBO J.* 9, 2361-2365.
- Hageman, T.C., Endres, G.F., & Scheraga, H.A. (1975). Mechanism of action of thrombin on fibrinogen; on the role of the A chain of bovine thrombin in specificity and differentiating between thrombin and trypsin. *Arch. Biochem. Biophys.* 171, 327-336.
- Hageman, T.C. & Scheraga, H.A. (1974). Mechanism of action of thrombin on fibrinogen; reaction of the N-terminal CNBR fragment from the A $\alpha$  chain of human fibrinogen with bovine thrombin. *Arch. Biochem. Biophys.* 164, 707-715.
- Hartley, B.S. & Kauffman, D. (1966). Corrections to the amino acid sequence of bovine chymotrypsinogen A. *Biochem. J.* 101, 229-231.
- Hartley, B.S. & Shotton, D.M. (1971). Pancreatic elastase. In *Enzymes*, Vol. 3 (Boyer, P.D., Ed.), pp. 323-373. Academic Press, New York, London.
- Haruyama, H. & Wüthrich, K. (1989). The conformation of recombinant disulfatohirudin in aqueous solution determined by nuclear magnetic resonance. *Biochemistry* 28, 4301-4312.
- Henriksen, R.A. & Mann, K.G. (1988). Identification of the primary structural defect in the dysfibrinogen Thrombin Quick I: Substitution of cysteine for arginine-382. *Biochemistry* 27, 9160-9165.
- Henriksen, R.A. & Mann, K.G. (1989). Substitution of valine for glycine-558 in the congenital dysfibrinogen Thrombin Quick II alters primary substrate specificity. *Biochemistry* 28, 2078-2082.
- Heuck, C.C., Schiele, V., Horn, D., Fronda, D., & Ritz, E. (1985). The role of surface charge on the accelerating action of heparin on the antithrombin III-inhibited activity of  $\alpha$ -thrombin. *J. Biol. Chem.* 260, 4598-4603.
- Hofsteenge, J., Braun, P.J., & Stone, S.R. (1988). Enzymatic properties of proteolytic derivatives of human  $\alpha$ -thrombin. *Biochemistry* 27, 2144-2151.
- Hofsteenge, J. & Stone, S.R. (1987). The effect of thrombomodulin on the cleavage of fibrinogen fragments by thrombin. *Eur. J. Biochem.* 168, 49-56.
- Hofsteenge, J., Taguchi, H., & Stone, S.R. (1986). Effect of thrombomodulin on the kinetics of the interaction of thrombin with substrates and inhibitors. *Biochem. J.* 237, 243-251.
- Hogg, D.H. & Blombäck, B. (1978). The mechanism of the fibrinogen-thrombin reaction. *Thromb. Res.* 12, 953-964.
- Hogg, P.J. & Jackson, C.M. (1989). Fibrin monomer protects thrombin from inactivation by heparin-antithrombin III: Implications for heparin efficacy. *Proc. Natl. Acad. Sci. USA* 86, 3619-3623.
- Horne, McD.K. & Gralnick, H.R. (1983). The oligosaccharide of human thrombin: Investigations of functional significance. *Blood* 63, 188-194.
- Huber, R. & Bode, W. (1978). Structural basis of the activation and action of trypsin. *Acc. Chem. Res.* 11, 114-122.
- Huber, R. & Carrell, R.W. (1989). Implications of the three-dimensional structure of  $\alpha_1$ -antitrypsin for structure and function of serpins. *Biochemistry* 28, 8951-8966.
- Huber, R., Kukla, D., Bode, W., Schwager, P., Bartels, K., Deisenhofer, J., & Steigemann, W. (1974). Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor. II. Crystallographic refinement at 1.9 Å resolution. *J. Mol. Biol.* 89, 73-101.
- Huber, R. & Schneider, M. (1985). A group refinement procedure in protein crystallography using Fourier transforms. *J. Appl. Crystallogr.* 18, 165-169.
- Jack, A. & Levitt, M. (1978). Refinement of large structures by simultaneous minimization of energy and R-factor. *Acta Crystallogr. A* 34, 931-935.
- Jackson, C.M. & Nemerson, Y. (1980). Blood coagulation. *Annu. Rev. Biochem.* 49, 765-811.
- Jakubowski, H.V. & Owen, W.G. (1989). Macromolecular specificity determinants on thrombin for fibrinogen and thrombomodulin. *J. Biol. Chem.* 264, 11117-11121.
- Jones, T.A. (1978). A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr.* 15, 23-31.
- Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: Pattern, recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22, 2577-2637.
- Kaczmarek, E. & McDonagh, J. (1988). Thrombin binding to the A $\alpha$ , B $\beta$ - and  $\gamma$ -chains of fibrinogen and to their remnants contained in fragment E. *J. Biol. Chem.* 263, 13896-13900.
- Kam, C.M., Fujikawa, W., & Powers, J.C. (1988). Mechanism-based isocoumarin inhibitors for trypsin and blood coagulation serine proteases: New anticoagulants. *Biochemistry* 27, 2547-2557.
- Kaminski, M. & McDonagh, J. (1987). Inhibited thrombins; interactions with fibrinogen and fibrin. *Biochem. J.* 242, 881-887.
- Karpatkin, S. & Karpatkin, M. (1974). Inhibition of the enzymatic activity of thrombin by concanavalin A. *Biochem. Biophys. Res. Commun.* 57, 1111-1118.
- Karshikov, A.D., Engh, R., Bode, W., & Atanasov, B.P. (1989). Electrostatic interactions in proteins: Calculations of the electrostatic term of free energy and the electrostatic potential field. *Eur. Biophys. J.* 17, 287-297.
- Kawabata, S., Morita, T., Iwanaga, S., & Igarashi, H. (1985). Staphylocoagulase-binding region in human prothrombin. *J. Biochem.* 97, 325-331.
- Kettner, C., Mersinger, L., & Knabb, R. (1991). The selective inhibition of thrombin by peptides of boronarginine. *J. Biol. Chem.* 265, 18289-18297.
- Kettner, C. & Shaw, E. (1979). D-Phe-Pro-Arg CH<sub>2</sub>Cl—A selective affinity label for thrombin. *Thromb. Res.* 14, 969-973.
- Kettner, C. & Shaw, E. (1981). Inactivation of trypsin-like enzymes with peptides of arginine chloromethyl ketone. *Methods Enzymol.* 80, 826-842.
- Klapper, I., Hagstrom, R., Fine, R., Sharp, K., & Honig, B. (1986). Focusing of electric fields in the active site of Cu-Zn superoxide dismutase: Effect of ionic strength and amino-acid modification. *Proteins Struct. Funct. Genet.* 1, 47-59.
- Landis, C., Köhler, R.A., & Fenton, J.W., II (1981). Human thrombins. *J. Biol. Chem.* 256, 4604-4610.
- Laskowski, M. & Kato, I. (1980). Protein inhibitors of proteinases. *Annu. Rev. Biochem.* 49, 593-626.
- LeBonniec, B.F. & Esmon, C.T. (1991). Glu-192  $\rightarrow$  Gln substitution in thrombin mimics the catalytic switch induced by thrombomodulin. *Proc. Natl. Acad. Sci. USA* 88, 7371-7375.
- Le Bonniec, B.F., MacGillivray, R.T.A., & Esmon, C.T. (1991). Thrombin Glu-39 restricts the P'3 specificity to nonacidic residues. *J. Biol. Chem.* 266, 13796-13803.
- Lee, B. & Richards, F.M. (1971). The interpretation of protein structure: Estimation of static accessibility. *J. Mol. Biol.* 55, 379-400.
- Levitt, M. (1974). Energy refinement of egg-white lysozyme. *J. Mol. Biol.* 82, 393-420.
- Lewis, S.D., Lorand, L., Fenton, J.W., II, & Shafer, J.A. (1987). Catalytic competence of human  $\alpha$ - and  $\gamma$ -thrombin in the activation of fibrinogen and factor XIII. *Biochemistry* 26, 7597-7603.
- Li, E.H.H., Fenton, J.W., II, & Feinman, R.D. (1976). The role of heparin in the thrombin-antithrombin III reaction. *Arch. Biochem. Biophys.* 175, 153-159.
- Liem, R.K.H. & Scheraga, H.A. (1974). Mechanism of action of thrombin on fibrinogen N. Further mapping of the active sites of thrombin and trypsin. The binding of thrombin and fibrin. *Arch. Biochem. Biophys.* 160, 333-339.
- Liu, C.Y., Nossel, H.L., & Kaplan, K.L. (1979). The binding of thrombin by fibrin. *J. Biol. Chem.* 254, 10421-10425.
- Liu, L.W., Vu, T.K.H., Esmon, C.T., & Coughlin, S.R. (1991). The region of the thrombin receptor resembling hirudin binds to thrombin and alters enzyme specificity. *J. Biol. Chem.* 266, 16977-16980.
- Löbermann, H., Tokuyama, R., Deisenhofer, J., & Huber, R. (1984). Human  $\alpha_1$ -proteinase inhibitor. Crystal structure analysis of two crystal modifications, molecular model and preliminary analysis of the implications for function. *J. Mol. Biol.* 177, 531-556.
- Lottenberg, R., Hall, J.A., Blinder, M., Binder, E.P., & Jackson, C.M. (1983). The action of thrombin on peptide *p*-nitroanilide substrates.

- Substrate selectivity and examination of hydrolysis under different reaction conditions. *Biochim. Biophys. Acta* 742, 539-557.
- Lundblad, R.L., Noyes, C.M., Featherstone, G.L., Harrison, J.H., & Jenzano, J.W. (1988). The reaction of bovine  $\alpha$ -thrombin with tetranitromethane. Characterization of the modified protein. *J. Biol. Chem.* 263, 3729-3734.
- Lundblad, R.L., Noyes, C.M., Mann, K.G., & Kingdon, H.S. (1979). The covalent differences between bovine  $\alpha$ - and  $\beta$ -thrombin; a structural explanation for the changes in catalytic activity. *J. Biol. Chem.* 254, 8524-8528.
- Luzzati, V. (1952). Traitement statistique des erreurs dans la détermination des structures cristallines. *Acta Crystallogr.* 5, 802-810.
- MacGillivray, R.T.A. & Davie, E.W. (1984). Characterization of bovine prothrombin mRNA and its translation product. *Biochemistry* 23, 1626-1634.
- Magnusson, S., Peterson, T.E., Sottrup-Jensen, L., & Claessens, H. (1975). Complete primary structure of prothrombin: Isolation, structure and reactivity of ten carboxylated glutamic acid residues and regulation of prothrombin activation by thrombin. In *Proteases and Biological Control* (Reich, E., Rifkin, D.B., & Shaw, E., Eds.), pp. 123-149. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Mao, S.J.T., Yates, M.T., Owen, T.J., & Krstenansky, J.L. (1988). Interaction of hirudin with thrombin: Identification of a minimal binding domain of hirudin that inhibits clotting activity. *Biochemistry* 27, 8170-8173.
- Maraganore, J.M., Bourdon, P., Jablonski, J., Ramachandran, K.L., & Fenton, J.W., II (1990). Design and characterization of hirulogs: A novel class of bivalent peptide inhibitors of thrombin. *Biochemistry* 29, 7095-7101.
- Marsh, H.C., Meinwald, Y.C., Lee, S., Martinelli, R.A., & Scheraga, H.A. (1985). Mechanism of action of thrombin on fibrinogen. Direct evidence for the involvement of phenylalanine at position P9. *Biochemistry* 24, 2806-2812.
- Martin, P.D., Robertson, W., Turk, D., Bode, W., & Edwards, B.F.P. (1992). The structure of residues 7-16 of the A $\alpha$ -chain of human fibrinogen bound to bovine thrombin at 2.3 Å resolution. *J. Biol. Chem.* (submitted).
- Matsuzaki, T., Sasaki, C., Okumura, C., & Umeyama, H. (1989). X-ray analysis of a thrombin inhibitor-trypsin complex. *J. Biochem. (Tokyo)* 105, 949-952.
- Matthew, J.B. (1985). Electrostatic effects in proteins. *Annu. Rev. Biophys. Biochem.* 14, 387-417.
- McGowan, E.B. & Detwiler, T.C. (1986). Modified platelet responses to thrombin. Evidence for two types of receptors or coupling mechanisms. *J. Biol. Chem.* 261, 739-746.
- Meinwald, Y.C., Martinelli, R.A., Van Nispen, J.W., & Scheraga, H.A. (1980). Mechanism of action of thrombin on fibrinogen. Size of the A $\alpha$  fibrinogen-like peptide that contacts the active site of thrombin. *Biochemistry* 19, 3820-3825.
- Meloun, B., Kluh, I., Kostka, V., Moravsek, L., Prusik, Z., Vanecek, J., Keil, B., & Sorm, F. (1966). Covalent structure of bovine chymotrypsin A. *Biochim. Biophys. Acta*, 130, 543-546.
- Messerschmidt, A. & Pflugrath, J.W. (1987). Crystal orientation and X-ray pattern prediction routines for area detector diffractometer systems in macromolecular crystallography. *J. Appl. Crystallogr.* 20, 306-315.
- Messerschmidt, A., Schneider, M., & Huber, R. (1990). ABSCOR: A scaling and absorption correction program for the FAST area detector diffractometer. *J. Appl. Crystallogr.* 23, 436-439.
- Meyer, E., Cole, G., Radhakrishnan, R., & Epp, O. (1988). Structure of native porcine pancreatic elastase at 1.65 Å resolution. *Acta Crystallogr. B* 44, 26-38.
- Mikes, O., Holeysovsky, V., Tomasek, V., & Sorm, F. (1966). Covalent structure of bovine trypsinogen. The position of the remaining amides. *Biochem. Biophys. Res. Commun.* 24, 346-352.
- Miyata, T., Morita, T., Inomoto, T., Kawauchi, S., Shirakami, A., & Iwanaga, S. (1987). Prothrombin Tokushima, a replacement of arginine-418 by tryptophan that impairs the fibrinogen clotting activity of derived thrombin Tokushima. *Biochemistry* 26, 1117-1122.
- Naski, M.C., Fenton, J.W., II, Maraganore, J.M., Olson, S.T., & Shafer, J.A. (1990). The COOH-terminal domain of hirudin. An exosite-directed competitive inhibitor of the action of  $\alpha$ -thrombin on fibrinogen. *J. Biol. Chem.* 265, 13484-13489.
- Needleman, S.B. & Wunsch, C.D. (1970). A general method applicable to the search for similarities in the amino acid sequences of two proteins. *J. Mol. Biol.* 48, 443-453.
- Nesheim, M.E. (1983). A simple rate law that describes the kinetics of the heparin-catalyzed reaction between antithrombin III and thrombin. *J. Biol. Chem.* 258, 14708-14717.
- Ni, F., Konishi, Y., Frazier, R.B., & Scheraga, H.A. (1989a). High-resolution NMR studies of fibrinogen-like peptides in solution: Interaction of thrombin with residues 1-23 of the A $\alpha$  chain of human fibrinogen. *Biochemistry* 28, 3082-3094.
- Ni, F., Meinwald, Y.C., Vasquez, M., & Scheraga, H.A. (1989b). Structure of a thrombin-bound peptide corresponding to residues 7-16 of the A $\alpha$ -chain of human fibrinogen. *Biochemistry* 28, 3094-3105.
- Nilsson, B., Horne, M.K., & Gralnick, H.R. (1983). The carbohydrate of human thrombin: Structural analysis of glycoprotein oligosaccharides by mass spectrometry. *Arch. Biochem. Biophys.* 224, 127-133.
- Noé, G., Hofsteenge, J., Rovelli, G., & Stone, S.R. (1988). The use of sequence specific antibodies to identify a secondary binding site in thrombin. *J. Biol. Chem.* 263, 11729-11735.
- Okamoto, S., Hijikata, A., Kikumoto, R., Tonomura, S., Hara, N., Ninomiya, K., Maruyama, A., Sugano, M., & Tamao, Y. (1981). Potent inhibition of thrombin by the newly synthesized arginine derivative no. 805. The importance of stereostructure of its hydrophobic carboxamide portion. *Biochem. Biophys. Res. Commun.* 101, 440-446.
- Olson, S.T. & Shore, J.D. (1982). Demonstration of a two-step reaction mechanism for inhibition of  $\alpha$ -thrombin by antithrombin III and identification of the step affected by heparin. *J. Biol. Chem.* 257, 14891-14895.
- Olson, T.A., Sonder, S.A., Wilner, G.D., & Fenton, J.W., II (1986). Heparin binding in proximity to the catalytic site of human  $\alpha$ -thrombin. *Ann. N.Y. Acad. Sci.* 485, 96-103.
- Patthy, L. (1985). Evolution of proteases of blood coagulation and fibrinolysis by assembly of modules. *Cell* 41, 657-663.
- Pflugrath, J.W., Saper, M.A., & Quiocho, F.A. (1984). In *Methods and Application in Crystallographic Computing* (Hall, S. & Ashiaka, T., Eds.), pp. 404-407. Clarendon Press, Oxford.
- Pomerantz, M.W. & Owen, W.G. (1978). A catalytic role for heparin. Evidence for a ternary complex of heparin cofactor thrombin and heparin. *Biochim. Biophys. Acta* 535, 66-77.
- Preissner, K.T., DeVos, V., & Müller-Berghaus, G. (1987). Binding of thrombin to thrombomodulin accelerates inhibition of the enzyme by antithrombin III. Evidence for a heparin-independent mechanism. *Biochemistry* 26, 2521-2528.
- Prescott, S.M., Seeger, A.R., Zimmerman, G.A., McIntyre, T.M., & Maraganore, J.M. (1990). Hirudin-based peptides block the inflammatory effects of thrombin on endothelial cells. *J. Biol. Chem.* 265, 9614-9616.
- Priestle, J.P. (1988). RIBBON: A stereo cartoon drawing program for proteins. *J. Appl. Crystallogr.* 21, 572-576.
- Ramakrishnan, C. & Ramachandran, G.N. (1965). Stereochemical criteria for polypeptide and protein chain conformations. II. Allowed conformations for a pair of peptide units. *Biophys. J.* 5, 909-933.
- Read, R.J. & James, M.N.G. (1986). Introduction to the proteinase inhibitors: X-ray crystallography. In *Proteinase Inhibitors* (Barret, A.J. & Salvesen, G., Eds.), pp. 301-336. Elsevier, Amsterdam.
- Remington, S.J., Woodbury, R.G., Reynolds, R.A., Matthews, B.W., & Neurath, H. (1988). The structure of rat mast cell protease II at 1.9 Å resolution. *Biochemistry* 27, 8097-8105.
- Richards, F.M. (1977). Areas, volumes, packing, and protein structure. *Annu. Rev. Biophys. Bioeng.* 6, 151-170.
- Richardson, J.S. & Richardson, D.C. (1990). Principles and patterns of protein conformation. In *Prediction of Protein Structure and the Principles of Protein Conformation* (Fasman, G., Ed.), pp. 1-98. Plenum Press, New York.
- Rosenberg, R.D. & Damus, P.S. (1973). The purification and mechanism of action of human antithrombin-heparin cofactor. *J. Biol. Chem.* 248, 6490-6505.
- Rossmann, M.G. & Argos, P. (1975). A comparison of the heme binding pocket in globins and cytochrome b5. *J. Biol. Chem.* 250, 7525-7532.
- Rydell, T.J., Ravichandran, K.G., Tulinsky, A., Bode, W., Huber, R., Roitsch, C., & Fenton, J.W., II (1990). The structure of a complex of recombinant hirudin and human  $\alpha$ -thrombin. *Science* 249, 277-280.



- Rydel, T.J., Tulinsky, A., Bode, W., & Huber, R. (1991). Refined structure of the hirudin-thrombin complex. *J. Mol. Biol.* 221, 583-601.
- Schechter, I. & Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* 27, 157-162.
- Scheraga, H.A. (1977). Active site mapping of thrombin. In *Chemistry and Biology of Thrombin* (Lundblad, R.L., Fenton, J.W., II, & Mann, K.G., Eds.), pp. 145-158. Ann Arbor Science Publ., Ann Arbor, Michigan.
- Schirmer, T., Bode, W., & Huber, R. (1987). Refined three-dimensional structure of two cyanobacterial C-phycoerythrins at 2.1 and 2.5 Å resolution. *J. Mol. Biol.* 196, 677-695.
- Schuman, M.A. (1986). Thrombin—Cellular interactions. *Ann. N.Y. Acad. Sci.* 485, 228-239.
- Skaug, K. & Christenson, T.B. (1971). The significance of the carbohydrate constituents of bovine thrombin for the clotting activity. *Biochim. Biophys. Acta* 230, 627-629.
- Sonder, S.A. & Fenton, J.W., II (1984). Proflavin binding within the fibrinopeptide groove adjacent to the catalytic site of human  $\alpha$ -thrombin. *Biochemistry* 23, 1818-1823.
- Steigemann, W. (1974). Die Entwicklung und Anwendung von Rechenverfahren und Rechenprogrammen zur Strukturanalyse von Proteinen am Beispiel des Trypsin-Trypsininhibitor Komplexes, des freien Inhibitors und der L-Asparaginase. Ph.D. Thesis, University of Munich.
- Stone, S.R., Braun, P.J., & Hofsteenge, J. (1987). Identification of regions of  $\alpha$ -thrombin involved in its interaction with hirudin. *Biochemistry* 26, 4617-4624.
- Stone, S.R. & Hofsteenge, J. (1986). Kinetics of the inhibition of thrombin by hirudin. *Biochemistry* 25, 4622-4628.
- Stubbs, M., Oschkinat, H., Mayr, I., Huber, R., Angliker, H., Stone, S.R., & Bode, W. (1992). The interaction of thrombin with fibrinogen—A structural basis for its specificity. *Eur. J. Biochem.* (in press).
- Stürzebecher, J., Markwardt, F., Voigt, B., Wagner, G., & Walsmann, P. (1983). Cyclic amides of  $N\alpha$ -arylsulfonylaminoacylated 4-amidinophenylalanine—Tight binding of thrombin. *Thromb. Res.* 29, 635-642.
- Sugawara, Y., Birktoft, J.J., & Berliner, L.J. (1986). Human  $\alpha$ - and  $\gamma$ -thrombin inhibition by trypsin inhibitors supports predictions from molecular graphics experiments. *Semin. Thromb. Hemostasis* 12, 209-210.
- Suzuki, K., Nishioka, J., & Hayashi, T. (1990). Localization of thrombomodulin-binding site within human thrombin. *J. Biol. Chem.* 265, 13263-13267.
- Svendsen, L., Blombäck, B., Blombäck, M., & Olsson, P.I. (1972). Synthetic chromogenic substrates for determination of trypsin, thrombin and thrombin-like enzymes. *Thromb. Res.* 1, 267-278.
- Tanford, D. & Kirkwood, J.G. (1957). Theory of protein titration curves. I. General equations for impenetrable spheres. *J. Am. Chem. Soc.* 79, 5333-5339.
- Thomas, K.A., Smith, G.M., Thomas, T.B., & Feldmann, R.J. (1982). Electronic distributions within protein phenylalanine aromatic rings are reflected by the three-dimensional oxygen atom environments. *Proc. Natl. Acad. Sci. USA* 79, 4843-4847.
- Thompson, A.R. (1976). High affinity binding of human and bovine thrombin to *p*-chlorobenzylamido- $\epsilon$ -aminocaproyl-agarose. *Biochim. Biophys. Acta* 422, 200-209.
- Tollefsen, D.M., Majerus, D.W., & Blank, M.K. (1982). Heparin cofactor II. Purification and properties of a heparin-dependent inhibitor of thrombin in human plasma. *J. Biol. Chem.* 257, 2162-2169.
- Toma, K. & Suzuki, K. (1989). Mapping active sites of blood coagulation serine proteinases—activated protein C and thrombin—on simple graphics models. *J. Mol. Graphics* 7, 146-149.
- Tsiang, M., Lentz, S.R., Dittman, W.A., Scarpato, E.M., & Sadler, J.E. (1990). Equilibrium binding of thrombin to recombinant human thrombomodulin: Effect of hirudin, fibrinogen, factor Va and peptide analogues. *Biochemistry* 29, 10602-10612.
- Tsukada, H. & Blow, D. (1985). Structure of  $\alpha$ -chymotrypsin refined at 1.68 Å resolution. *J. Mol. Biol.* 184, 703-711.
- Turk, D., Stürzebecher, J., & Bode, W. (1991). Geometry of binding of  $\alpha$ -tosylated piperidides of *m*-amidino, *p*-amidino- and *p*-guanidino phenylalanine to thrombin and trypsin. X-ray crystal structures of their trypsin complexes and modeling of their thrombin complexes. *FEBS Lett.* 287, 133-138.
- Vali, Z. & Scheraga, H.A. (1988). Localization of the binding site on fibrin for the secondary binding site of thrombin. *Biochemistry* 27, 1956-1963.
- Van Nispen, J.W., Hageman, T.C., & Scheraga, H.A. (1977). Mechanism of action of thrombin on fibrinogen; the reaction of thrombin with fibrinogen-like peptides containing 11, 14 and 16 residues. *Arch. Biochem. Biophys.* 182, 227-243.
- Vu, T.K.H., Hung, D.T., Wheaton, V.I., & Coughlin, S.R. (1991). Molecular cloning of a functional thrombin receptor reveals a novel proteolytic mechanism of receptor activation. *Cell* 64, 1057-1068.
- Wallace, A., Rovelli, G., Hofsteenge, J., & Stone, S.R. (1989). Effect of heparin on the glia-derived nexin-thrombin interaction. *Biochem. J.* 257, 191-196.
- Walsh, K.A. & Neurath, H. (1964). Trypsinogen and chymotrypsinogen as homologous proteins. *Proc. Natl. Acad. Sci. USA* 52, 884-889.
- Walsmann, P. & Markwardt, F. (1981). Biochemische und pharmakologische Aspekte des Thrombininhibitors Hirudin. *Pharmazie* 36, 653-660.
- Wang, D., Bode, W., & Huber, R. (1985). Bovine chymotrypsinogen A. X-ray crystal structure analysis & refinement of a new crystal form at 1.8 resolution. *J. Mol. Biol.* 185, 595-624.
- Wei, A.-Z., Mayr, I., & Bode, W. (1988). The refined 2.3 Å crystal structure of human leukocyte elastase in a complex with a valine chloromethylketone inhibitor. *FEBS Lett.* 234, 367-373.
- White, G.C., Lundblad, R.L., & Griffith, M.J. (1981). Structure-function relations in platelet-thrombin reactions. *J. Biol. Chem.* 256, 1763-1766.
- Wu, Q., Sheehan, J.P., Tsiang, M., Lentz, S.R., Birktoft, J.J., & Sadler, J.E. (1991). Single amino acid substitutions dissociate fibrinogen-clotting and thrombomodulin-binding activities of human thrombin.

# Protein Engineering Thrombin for Optimal Specificity and Potency of Anticoagulant Activity *in Vivo*

M. Tsiang, L. R. Paborsky, W.-X. Li, A. K. Jain, C. T. Mao, K. E. Dunn, D. W. Lee, S. Y. Matsumura, M. D. Matteucci, S. E. Coutré,<sup>‡</sup> L. L. K. Leung,<sup>‡</sup> and C. S. Gibbs\*

Gilead Sciences Inc., 353 Lakeside Drive, Foster City, California 94404

Received July 3, 1996; Revised Manuscript Received September 16, 1996<sup>©</sup>

**ABSTRACT:** Previous alanine scanning mutagenesis of thrombin revealed that substitution of residues W50, K52, E229, and R233 (W60d, K60f, E217, and R221 in chymotrypsinogen numbering) with alanine altered the substrate specificity of thrombin to favor the anticoagulant substrate protein C. Saturation mutagenesis, in which residues W50, K52, E229, and R233 were each substituted with all 19 naturally occurring amino acids, resulted in the identification of a single mutation, E229K, that shifted the substrate specificity of thrombin by 130-fold to favor the activation of the anticoagulant substrate protein C over the procoagulant substrate fibrinogen. E229K thrombin was also less effective in activating platelets (18-fold), was resistant to inhibition by antithrombin III (33-fold and 22-fold in the presence and absence of heparin), and displayed a prolonged half-life in plasma *in vitro* (26-fold). Thus E229K thrombin displayed an optimal phenotype to function as a potent and specific activator of endogenous protein C and as an anticoagulant *in vivo*. Upon infusion in Cynomolgus monkeys E229K thrombin caused an anticoagulant effect through the activation of endogenous protein C without coincidentally stimulating fibrinogen clotting and platelet activation as observed with wild-type thrombin. In addition, E229K thrombin displayed enhanced potency *in vivo* relative to the prototype protein C activator E229A thrombin. This enhanced potency may be attributable to decreased clearance by antithrombin III, the principal physiological inhibitor of thrombin.

Activation of the multienzyme blood coagulation cascade results in the conversion of the inactive zymogen prothrombin into active thrombin (Rosenberg, 1987; Jackson & Nemerson, 1980). Thrombin is a multifunctional serine protease that recognizes and proteolytically activates multiple substrates (Mann & Lundblad, 1987; Esmon 1987) that have both procoagulant and anticoagulant functions. Through precise regulation of its substrate specificity thrombin may function as a promoter of blood coagulation, mediating both fibrin clot formation and platelet aggregation, or alternatively thrombin may act as a regulator of coagulation by activating a natural anticoagulant mechanism, the protein C (PC) pathway.

At sites of vascular injury or inflammation, thrombin recognizes two substrates with key procoagulant functions. The hydrolysis of fibrinogen results in the formation of an insoluble fibrin clot, and the cleavage of the platelet thrombin receptor results in the generation of a tethered ligand that promotes platelet activation and aggregation (Vu et al., 1991). Additional procoagulant substrates include factor XIII (Lorand & Radek, 1992), which stabilizes the clot by cross-linking fibrin monomers, and factors V, VIII (Mann et al., 1988), and XI (Gailani & Broze, 1991) that perpetuate the clotting stimulus and the generation of new thrombin through feedback activation of the coagulation cascade. Thrombin that escapes the site of an injury becomes bound to the cofactor thrombomodulin (TM), a transmembrane protein expressed on endothelial cells that line the vasculature. Upon

binding to TM the activity of thrombin toward the procoagulant substrates is blocked, and a proposed Ca<sup>2+</sup>-dependent conformational change enhances the recognition and activation of the anticoagulant substrate PC. Activated protein C (aPC) cleaves and inactivates activated coagulation factors Va and VIIIa to repress the coagulation cascade and may also enhance fibrinolysis by inactivating PAI-1, the principal physiological inhibitor of tissue plasminogen activator (Comp & Esmon, 1981; van Hinsbergh, 1985), or indirectly by inhibiting the activation of TAFI (thrombin activatable inhibitor of thrombolysis) (Bajzar et al., 1995). Thus the aPC pathway may serve as a mechanism to attenuate the extent of blood clotting and prevent propagation of the clot beyond the site of an injury (Esmon, 1987).

Several approaches have been used to exploit the PC pathway as a means to achieve anticoagulation *in vivo* including infusion of soluble thrombomodulin (Gomi et al., 1990), infusion of a PC-activating protease isolated from the venom of snakes belonging to the genus *Akistrodon* (Strukova et al., 1989; Kogan et al., 1993), direct infusion of exogenously activated PC (Comp & Esmon, 1981; Gruber et al., 1989, 1990; Okajima et al., 1990; Taylor et al., 1987), or infusion of a variant of PC that is more susceptible to activation by thrombin (Richardson et al., 1992; Kurz et al., 1994).

An alternative approach has been to take advantage of enzyme turnover and activate endogenous PC by infusion of thrombin itself. Infusion of thrombin in dogs resulted in systemic anticoagulation (Comp et al., 1982), and infusion of thrombin in baboons was shown to be effective in reducing thrombus formation in arteriovenous shunts (Hanson et al., 1993). In both cases, coinfusion of antibodies that blocked PC also blocked the anticoagulant effect of the thrombin

\* To whom correspondence should be addressed: telephone (415) 573-4852; Fax (415) 573-4890.

<sup>‡</sup> Present address: Division of Hematology, Stanford University School of Medicine, Stanford, CA 94305-5112.

<sup>©</sup> Abstract published in *Advance ACS Abstracts*, December 1, 1996.



infusion, strongly implicating the aPC anticoagulant pathway. However, the effects of the procoagulant activities of thrombin were also evident particularly at higher doses. Elevation of markers of fibrin formation and platelet activation and consumption of fibrinogen and clotting factors were indicative of intravascular coagulation.

One approach to minimizing the procoagulant functions of thrombin while retaining the anticoagulant effect has been to infuse a reversibly acylated form of thrombin [(guanidinobenzoyl)thrombin] that was maintained in inactive form prior to thrombomodulin binding (McBane et al., 1995). However, preliminary studies demonstrating that the PC activation and fibrinogen clotting activities of thrombin could be dissected by point mutations (Wu et al., 1991) suggested to us that a genetically engineered variant of human thrombin with impaired specificity for the procoagulant substrates that retained anticoagulant activity might be a more attractive anticoagulant agent.

A library of 62 thrombin mutants, in which the charged and polar amino residues that were highly exposed on the surface of the molecule were substituted with the small neutral amino acid alanine, was screened for mutants that were defective in fibrinogen clotting yet retained their ability to activate PC in the presence of TM. Four mutants (W50A, K52A, E229A, and R233A) displayed significant shifts in their substrate specificity that favored the anticoagulant substrate, PC (Tsiang et al., 1995). The E229A mutant, which displayed a 22-fold switch in substrate specificity, was selected as a prototype and was demonstrated to function as a selective endogenous protein C activator (PCA) in Cynomolgus monkeys, causing reversible anticoagulation without activating platelets or fibrin clot formation (Gibbs et al., 1995).

In this study, we attempted to optimize the specificity of thrombin toward PC by individually substituting thrombin residues W50, K52, E229, and R233 with all 19 naturally occurring amino acids. Screening of this collection of all 76 possible substitutions identified a mutant, E229K, with enhanced specificity for PC and enhanced anticoagulant potency *in vivo*.

## EXPERIMENTAL PROCEDURES

**Materials.** Reacti-BIND maleic anhydride activated polystyrene 96-well plates were from Pierce (catalog no. 15110). PPACK (D-Phe-Pro-Arg chloromethyl ketone) was from Sigma. Purified human  $\alpha$ -thrombin and protein C (free of detectable prothrombin) were from Haematologic Technologies. Purified human antithrombin III was from Enzyme Research Laboratories. The anti-human prothrombin polyclonal antibody was from DAKO (catalog no. 325), and the goat anti-rabbit antibody conjugated with horseradish peroxidase was from Vector Laboratories. *Echis carinatus* snake venom was from Sigma. DEAE-Sepharose Fast Flow was from Pharmacia. Amberlite CG50 cation-exchange resin was from ICN Biomedicals.

**Construction, Transient Expression, and Screening of Thrombin Mutants.** The detailed methods describing the expression, processing, and assay of transiently expressed thrombin mutants have been described previously (Tsiang et al., 1995).

Thrombin mutants were constructed by oligonucleotide-directed mutagenesis as described previously (Gibbs & Zoller, 1991) using four synthetic oligonucleotide primers

where the codons encoding thrombin residues W50, K52, E229, and R233 were replaced with randomized codons NNS (where N = A or C or G or T and S = C or G). On average, clones encoding 13/19 of all possible amino acid substitutions at each position were identified following the sequencing of 48 clones. The remaining mutations were constructed using specific mutagenesis primers. The collection of 76 thrombin mutants was transiently expressed in COS-7 cells. Conditioned medium was concentrated by ultrafiltration and treated with *E. carinatus* venom to process prothrombin to thrombin. Processing was determined to be complete following analysis by Western blotting of reducing SDS-PAGE gels. Thrombin concentration in concentrated conditioned medium was determined by ELISA (see below). Specific amidolytic activity [S-2238 (H-D-Phe-Pip-Arg-pNA) hydrolysis], fibrinogen clotting activity, and PC activation in the presence of TM were determined as described previously (Tsiang et al., 1995). Specific fibrinogen clotting activity (FC) (expressed as a percent of wild-type thrombin activity) was determined from the time to clot formation following the addition of purified fibrinogen (to 0.4 mg/mL) to a predetermined amount of each thrombin mutant. Clotting activity was related to the clotting activity of wild-type thrombin using a standard curve of clotting time relative to thrombin concentration. Specific activity of protein C activation (PA) (expressed as a percent of wild-type thrombin activity) was determined by incubating each thrombin mutant with 5 nM TM and 887 nM PC. The concentration of aPC generated was determined from the hydrolysis of the chromogenic substrate S2366 following termination of the reaction by the addition of heparin/antithrombin III. Specific activity of PC activation was based on the concentration of thrombin/TM complexes in the reaction. The relative specificity for protein C activation over fibrinogen clotting was determined from the PA/FC ratio.

**Thrombin ELISA.** The expression level of each thrombin mutant was estimated using an immunoassay employing PPACK to capture and immobilize thrombin on a microtiter plate for subsequent detection using an anti-human prothrombin/thrombin polyclonal antibody. Purified human  $\alpha$ -thrombin was used to construct a standard curve for each assay. Microtiter plates were prepared freshly for each assay as follows: PPACK (15  $\mu$ M) in PBS was coated onto the bottom of a maleic anhydride activated 96-well plate and incubated overnight at room temperature. The PPACK was removed, and the excess sites on the plate were blocked for 30 min with 3% BSA in PBS and 0.05% Tween-20 (Tw). The blocking step was repeated, and the plate was then washed with wash buffer (PBS, 0.05% Tw). Triplicate aliquots of purified thrombin standards (0.5–12 ng/mL) and duplicate aliquots of thrombin mutants were diluted in dilution buffer (0.1% BSA in PBS 0.05% Tw) and then added to the plate and incubated for 2 h at room temperature. Thrombin was detected by incubating the plate for 1 h at room temperature with the rabbit anti-human prothrombin/thrombin polyclonal antibody (1:1000) diluted in dilution buffer followed by 1 h incubation at room temperature with a goat anti-rabbit antibody conjugated with horseradish peroxidase (1:2000) diluted in PBS and 0.1% Tw. The plate was washed between each incubation step and developed using the peroxidase substrate ABTS (Vector Laboratories). The absorbance at 405 nm was determined using a Molecular Devices plate reader. The concentration of each thrombin mutant was determined from the standard curve using data

from at least two dilutions each measured in duplicate. The assay was sensitive to 0.5 ng/mL thrombin, and the linear range extended beyond 12 ng/mL.

**Expression of Selected Thrombin Mutants Using an Expression Vector Based on the Sindbis Virus.** An *Xba*I fragment encoding the entire prothrombin coding region was isolated from BS(KS)-hFII (obtained from Ross MacGillivray, University of British Columbia) (Friezner Degen et al., 1983) and was cloned into the *Xba*I site downstream of the subgenomic RNA promoter in the Sindbus virus-based expression vector SINrep5 (Bredenbeek et al., 1993) to generate SINrep5-hPT. A 1301 bp *Kas*I–*Apal* fragment that spans the region encoding the thrombin B-chain was isolated from pRc/CMV-hPT clones expressing selected mutants (E229A, E229K, E229S, E229W, E229Y, R233F, W50E, W50I) and used to replace the corresponding *Kas*I–*Apal* fragment in SINrep5-hPT, allowing expression of thrombin mutants from the Sindbis subgenomic promoter.

**Cell Culture and Cell Lines.** COS-7 and BHK-21 cells were grown and passaged in media I (DMEM containing 4.5 mg/mL glucose, 3.7 mg/mL NaHCO<sub>3</sub>, 10% CS, 2 mM glutamine, 100 units/mL penicillin, and 100 µg/mL streptomycin). During the prothrombin secretion stage, BHK-21 cells were incubated with media II (serum-free media I supplemented with 10 µg/mL vitamin K, 5 µg/mL insulin, 5 µg/mL transferrin, and 5 µg/mL fetuin). The replicon packaging helper cell line BHK-BBneo was a generous gift from Dr. Sondra Schlesinger (Washington University, St. Louis, MO). This cell line was derived by stable transfection of BHK-21 cells (Bredenbeek et al., 1993) with a construct comprised of the Sindbis helper RNA (DH-BB) encoding the Sindbis structural proteins (Liljeström & Garoff, 1991; Liljeström et al., 1991) followed by a Neo<sup>R</sup> gene. In the absence of Sindbis replication machinery (supplied in *trans* through transfection of replicon RNAs), these cells do not express Sindbis-specific polypeptides. BHK-BBneo cells were passaged in media III [ $\alpha$ -MEM containing 10% FCS, 100 units/mL penicillin, 100 µg/mL streptomycin, and 800 µg/mL G418 (Gibco)].

**Generation of Sindbis Replicons and Expression of Prothrombin Mutants.** SINrep5-hPT mutants were linearized with *Not*I. The linearized plasmids (1–3 µg/50 µL of reaction) were used as template for the synthesis of replicon RNA by SP6 RNA polymerase in the presence of 1 mM 5' cap analog m<sup>7</sup>G(5')ppp(5')G (New England Biolabs) as previously described (Rice et al., 1987). BHK-BBneo cells were transfected with 2 pmol of replicon RNA by electroporation, diluted in 13.5 mL of media III, and plated in two T25 flasks (Liljeström et al., 1991). The media containing packaged replicons were harvested 48 h postelectroporation when the cytopathic effect was obvious. At this time, titers of packaged replicons were  $3 \times 10^8$ – $2 \times 10^9$  infectious units/mL (Schlesinger, personal communication). We assumed that all our P0 harvests had a titer of  $\sim 1 \times 10^8$  infectious units/mL. To amplify the replicon stock, new monolayers of BHK-BBneo were infected with the P0 harvest at an moi of  $\sim 1.0$ . The amplified replicon stock harvested 48 h postinfection was used to infect BHK-21 cells in roller bottles at an moi of  $\sim 6.0$ . Seventy two hours postinfection, conditioned medium containing secreted prothrombin was harvested. The expression level was  $\sim 4.6$  µg of prothrombin/10<sup>6</sup> cells or  $\sim 417$  µg of prothrombin/roller bottle.

**Purification of Thrombin Mutants.** The purification procedures for prothrombin and thrombin were modified from previously described protocols (Wu et al., 1991; Fenton et al. 1977). For each mutant, the conditioned medium from four roller bottles was clarified by centrifugation at 2500 rpm and vacuum filtration through Whatman No. 3 and Whatman No. 1 filter paper. Sindbis replicon particles in the media were inactivated by adding Triton X-100 to a final concentration of 0.2%. Benzamidine hydrochloride (5 mM) and barium sulfate (40 mg/mL) were added, and the medium was incubated on ice for 30 min. The medium was then centrifuged at 3000 rpm for 15 min, and the pellet was washed three times with 5 mM sodium acetate, pH 7.0. The prothrombin was eluted with 200 mM sodium citrate, pH 7.0, and 5 mM benzamidine hydrochloride and dialyzed against 0.1 M potassium phosphate, pH 7.5, containing 5 mM benzamidine hydrochloride with the final dialysis step performed in the absence of benzamidine hydrochloride. The dialysate was then loaded onto a DEAE-Sepharose Fast Flow column (10 mL) equilibrated in 0.1 M potassium phosphate, pH 7.5. The column was washed with the equilibration buffer, and the prothrombin was eluted with a 0.1–0.7 M potassium phosphate, pH 7.5, gradient. A sample from each fraction was activated by *E. carinatus* venom and assayed for amidolytic activity. The peak fractions were pooled and dialyzed against 0.02 M HEPES and 0.1 M NaCl, pH 8.0. The dialysate was then concentrated using a stirred cell with a PM30 membrane (Amicon). The prothrombin was processed to thrombin by *E. carinatus* venom that had been cleared of components that adsorb to Amberlite CG50 cation-exchange resin. Following activation, the resulting thrombin was immediately loaded onto an Amberlite CG50 column (8 mL) equilibrated in 0.02 M HEPES and 0.1 M NaCl, pH 8.0. The column was washed with equilibration buffer and eluted with a 0.10–1.0 M NaCl gradient containing 20 mM HEPES, pH 8.0. The fractions were assayed for amidolytic activity, pooled, and concentrated. The purity of each mutant was assessed by SDS-PAGE and silver staining, and the thrombin concentration was determined by thrombin ELISA.

**Kinetic Analysis of Fibrinopeptide A Release.** Fibrinopeptide A (FPA) release was carried out as previously described (Higgins et al., 1983). Briefly, 500 µL reactions of fibrinogen and mutant thrombin in reaction buffer containing 137 mM NaCl, 2.5 mM KCl, 0.9 mM CaCl<sub>2</sub>, 0.5 mM MgCl<sub>2</sub>, 9.47 mM sodium phosphate, pH 7.4, and 0.001% PEG-6000 were incubated at 37 °C and stopped with 43 µL of 3 M HClO<sub>4</sub>. The precipitated protein was spun down in a microcentrifuge at 15 000 rpm for 5 min. The supernatant was transferred to an automatic sample injector for injection of 250 µL into the HPLC system. FPA and FPB were separated on a C<sub>18</sub> column using a 15 min gradient of 22.5%–49.5% CH<sub>3</sub>CN in 0.1% TFA. FPA and FPB eluted at 30.6% and 32.4% CH<sub>3</sub>CN, respectively. For determination of EC<sub>50</sub> (the concentration of thrombin causing the release of 50% of FPA after 60 min at 37 °C), mutant thrombin concentration was varied from 0.5 to 30 nM, and the reactions were incubated for 60 min at 37 °C and stopped. The data were fitted to the empirical function  $P = P_{\max} (1 - e^{-(\ln 2)[\text{IIa}]/\text{EC}_{50}})$ , which best fit the data, using nonlinear regression to calculate the EC<sub>50</sub> and associated standard error.  $P$  is the concentration of FPA released with a thrombin concentration of [IIa], and  $P_{\max}$  is the maximum concentration of FPA that can be released.



For determination of the catalytic efficiency ( $k_{\text{cat}}/K_m$ ) of FPA release, a thrombin concentration that allowed the complete release of both FPA and FPB in 60 min was used. The kinetic parameters were determined by the analysis of differential progress curves (Graycar & Estell, 1987) generated at two fibrinogen concentrations, 1.4 and 21  $\mu\text{M}$ , from 0 to 60 min and from 0 to 70 min, respectively.

Briefly, for the simple one-substrate Michaelis–Menten mechanism with product inhibition, the differential rate equation is

$$v = \frac{dP}{dt} = \frac{V_m S}{K_m \left( 1 + \frac{(S_0 - S)}{K_i} \right) + S} \quad (1)$$

where  $S$  is the substrate concentration at time  $t$ ,  $S_0$  the substrate concentration at time  $t = 0$ ,  $(S_0 - S)$  the product concentration, and  $K_i$  the product inhibition binding constant. Equation 1 can be rearranged to give a linear plot of  $1/v$  versus  $1/S$ :

$$\frac{1}{v} = \frac{K_m(1 + (S_0/K_i))}{V_m} \frac{1}{S} + \frac{1 - (K_m/K_i)}{V_m} \quad (2)$$

In the absence of product inhibition,  $K_i$  approaches infinity and eq 2 simplifies to the Lineweaver–Burk equation. Performing a second progress curve using the same total enzyme but a new initial substrate level such that  $S_0' = aS_0$  allows for the  $K_i$  to be determined. The ratio of the slopes from eq 2 for each reaction gives

$$\frac{m'}{m} = \frac{1 + aS_0/K_i}{1 + S_0/K_i} \quad (3)$$

Rearranging eq 3 we have

$$K_i = \frac{(a - m'/m)S_0}{(m'/m - 1)} \quad (4)$$

Using the derived  $K_i$  and the slope  $m$  and  $y$ -intercept  $b$  from one of the two progress curves, the  $K_m$  and  $V_m$  are calculated as follows:

$$K_m = \frac{mK_i}{b(K_i + S_0 + m/b)} \quad (5)$$

$$V_m = \frac{K_m(1 + S_0/K_i)}{m} \quad (6)$$

The progress curve of FPA release can be expressed as the integrated form of the Michaelis–Menten equation:

$$t = k_1 \ln \left( \frac{[A\alpha]_0}{[A\alpha]_0 - [FPA]} \right) + k_2 [FPA] \quad (7)$$

where  $[FPA]$  is the concentration of FPA at time  $t$ ,  $[A\alpha]_0$  is the initial concentration of  $A\alpha$ -chains of fibrinogen,  $k_1$  is a constant, function of  $[A\alpha]_0$ ,  $V_m$ ,  $K_m$ , and  $K_i$  (the product inhibition constant), and  $k_2$  is another constant, function of  $V_m$ ,  $K_m$ , and  $K_i$ .  $[A\alpha]_0$ ,  $k_1$ , and  $k_2$  were determined by fitting

eq 7 to the data. The inverse rate  $1/v$  of FPA release at each time point was then calculated using the equation:

$$\frac{1}{v} = \frac{dt}{d[FPA]} = \frac{k_1}{([A\alpha]_0 - [FPA])} + k_2 \quad (8)$$

$1/v$  was plotted versus  $1/[A\alpha] = 1/([A\alpha]_0 - [FPA])$ , and the slope  $m$  and  $y$ -intercept  $b$  were then determined by curve fitting to a line.  $K_i$ ,  $K_m$ , and  $V_{\text{max}}$  were subsequently calculated according to eqs 4–6.

**Kinetic Analysis of Protein C Activation.** Cell lysates containing recombinant human TM were prepared as previously described (Tsiang et al., 1995). The concentration of thrombomodulin was calculated by direct equilibrium binding of  $^{125}\text{I}$ -DIP-thrombin to cells (Tsiang et al., 1990). Activation of PC was carried out in 50 mM Tris-HCl, pH 8.0, 2 mM  $\text{CaCl}_2$ , 100 mM NaCl, and 0.1% BSA. The aPC generated was assayed by hydrolysis of chromogenic substrate S-2366. To determine the affinity of mutant thrombin for TM, TM and PC were used at 0.5 nM and 1.0  $\mu\text{M}$ , respectively, with increasing concentrations of mutant thrombin (0.1–20 nM). The  $K_d$  for TM binding and the rate of PC activation at saturating thrombin concentration ( $V_s$ ) were determined by fitting the data to the binding equation  $[EL] = [E_0][L]/(K_d + [L])$ , where  $[EL]$  is the concentration of the thrombin/TM complex measured as the rate of PC activation and  $[E_0]$  is the concentration of TM measured as  $V_s$ : the rate of PC activation when TM is saturated with thrombin and the concentration of PC is 1  $\mu\text{M}$ . The background rate of PC activation in the absence of TM was subtracted.

To determine the catalytic efficiency ( $k_{\text{cat}}/K_m$ ) of PC activation, steady-state kinetic conditions were used, with TM at 0.5 nM, mutant thrombin at the saturating concentration of 30 nM, and seven concentrations of PC (0.50, 0.66, 1.00, 1.30, 2.00, 4.00, and 10.00  $\mu\text{M}$ ). The initial velocity of PC activation for each concentration of PC was determined from a time course with time points taken at 7-min intervals. The  $K_m$  and  $k_{\text{cat}}$  for PC activation were again determined by fitting the Michaelis–Menten equation to the data assuming that the enzyme concentration was 0.5 nM, equal to the concentration of TM.

**Platelet Aggregation.** Platelet aggregation by recombinant wild-type and mutant thrombins (WT, E229A, E229K, and E229W) was performed with fresh citrated human platelet-rich plasma (PRP) using a Chrono-Log dual channel aggregometer (model 560-VS, Chrono-Log Corp., Havertown, PA). Recombinant human thrombin (WT or mutants) was added to 400  $\mu\text{L}$  of prewarmed PRP (37  $^\circ\text{C}$  for 2 min) to a final concentration of 0.19–1.29 nM (WT), 2.7–22.0 nM (E229A), 5.9–29.7 nM (E229K), or 12.9–49.7 nM (E229W). The platelet aggregation stimulated by addition of wild-type or mutant thrombins was monitored by the aggregometer for 5 min as an increase in light transmission. The extent of platelet aggregation was quantitated by measuring the area under the tracing of the aggregation curve from 0 to 1 min after the addition of thrombin. The extent of aggregation was expressed in the arbitrary unit of  $\text{cm}^2$  and plotted against the concentration of the agonist. The data were fitted to the empirical function which best fit the data:  $A = A_{\text{max}}/(1 + e^{n(\text{EC}_{50} - [I\text{Ia}] )})$ , where  $A$  is the extent of platelet aggregation at a thrombin concentration of  $[I\text{Ia}]$ ,  $A_{\text{max}}$  is the maximum extent of platelet aggregation, and  $n$  is a parameter that controls the steepness of the logistic curve. The  $\text{EC}_{50}$  value (concentration of thrombin required to induce half-maximal

aggregation after 1 min) and associated standard error were determined by nonlinear regression.

**Inhibition of Selected Thrombin Mutants by Antithrombin III in the Absence of Heparin.** Selected purified thrombin mutants were screened for resistance to ATIII by determining the concentration required for half-maximal inhibition ( $IC_{50}$ ) over a 30 min time period in the absence of heparin. Wild-type or mutant thrombins (5–80 nM) were incubated in selection buffer (20 mM Tris–acetate, pH 7.5, 140 mM NaCl, 5 mM KCl, 1 mM  $MgCl_2$ , 1 mM  $CaCl_2$ ) plus 0.1% BSA with increasing concentrations (0.008–8.0  $\mu M$ ) of purified ATIII at 25 °C. After 30 min the inhibition reaction was terminated by the addition of 100  $\mu M$  chromogenic peptide substrate S-2238. The residual activity due to uninhibited thrombin was determined by monitoring the hydrolysis of S-2238.  $IC_{50}$  values (ATIII concentration required for 50% inhibition of thrombin after 30 min) and the associated standard errors were determined by fitting the data to the empirical function:  $A = 100(IC_{50})^n / ((IC_{50})^n + [ATIII]^n)$ , which best fit the data, using nonlinear regression, where  $A$  is the percent thrombin activity remaining at an antithrombin III concentration of  $[ATIII]$  and  $n$  is a parameter that controls the steepness of the curve.

Second-order rate constants for inhibition of purified thrombin mutants were determined as described previously (Ye et al., 1994). Thrombin (20–60 nM) was incubated in selection buffer plus 0.1% BSA with increasing concentrations (0.2–6.0  $\mu M$ ) of purified ATIII at 25 °C. At various time intervals (1–95 min) aliquots of the reaction mixture were diluted 10-fold into 100  $\mu M$  S-2238 to terminate the inhibition reaction. The residual activity due to uninhibited thrombin was determined by monitoring the rate of hydrolysis of S-2238. Under pseudo-first-order conditions the time course of inhibition followed eq 9:

$$E_t = E_0 e^{-k't} \quad (9)$$

where  $E_0$  is the thrombin activity at time  $t = 0$  in  $A_{405}/s$ ,  $E_t$  the thrombin activity at time  $t$ , and  $k'$  the observed pseudo-first-order rate constant. The pseudo-first-order rate constant of the decrease in activity due to inhibition by ATIII ( $k'$ ) was determined by nonlinear regression, and the second-order rate constant of inhibition ( $k_2$ ) was determined from the slope of plots of  $k'$  versus ATIII concentration.

**Heparin-Dependent Inhibition by Antithrombin III.** The reaction buffer contained 20 mM Tris–acetate, pH 7.5, 140 mM NaCl, 5 mM KCl, 1 mM  $MgCl_2$ , 1 mM  $CaCl_2$ , and 0.1% PEG-8000. Inhibition by ATIII in the presence of heparin was performed in the presence of competing substrate (S-2238) as previously described (Sheehan et al., 1993). Briefly, 150 nM ATIII, 1.0 unit/mL heparin, and 150  $\mu M$  S-2238 were premixed in 590  $\mu L$  of reaction buffer. This mixture was added to a spectrophotometric cuvette containing 10  $\mu L$  of 300 nM mutant thrombin (final concentration of 5 nM), and hydrolysis of S-2238 was immediately measured for 2 min. The pseudo-first-order rate constant of inhibition ( $k'$ ) was determined by fitting eq 9 to the first derivative of the time course of S-2238 hydrolysis, which was directly calculated from the data. To determine the second-order rate constant  $k_2$ ,  $k'$  was corrected for substrate competition and divided by the inhibitor concentration  $[I]$  as shown in eq 10:

$$k_2 = \frac{k'(1 + [S]/K_m)}{[I]} \quad (10)$$

where  $[S]$  is the substrate concentration during the reaction and  $K_m$  the Michaelis constant of the substrate for each mutant thrombin.

**Clearance of Thrombin Activity from Human Plasma in Vitro.** Thrombin and thrombin mutants (100 nM) were mixed with 10 volumes of fibrinogen-deficient human plasma (George King Biomedical), and incubated at 25 °C. At various time intervals aliquots of the reaction mixture were diluted 50-fold into 100  $\mu M$  S-2238 to terminate the inhibition reaction. The residual activity due to uninhibited thrombin was determined by monitoring the hydrolysis of S-2238. The kinetics of the clearance of thrombin activity versus time approximately corresponded to a first-order exponential decay (eq 9). The half-life of thrombin and thrombin variants in human plasma was estimated using nonlinear regression analysis.

**Anticoagulation Studies in Cynomolgus Monkeys.** A cell-line stably expressing E229K prothrombin was constructed by transfecting CHO cell-line AA8 with pRc/CMV-hPT containing the E229K mutation. E229K prothrombin was secreted into CHO serum-free medium (Gibco), harvested, processed to thrombin, and purified as described previously (Gibbs et al., 1995).

All animals received humane care consistent with published NIH guidelines (NIH, 1985). Four adult Cynomolgus monkeys received 10 min intravenous infusions of purified E229K at a dose of 0.5  $\mu g\ kg^{-1}\ min^{-1}$ , formulated in a total volume of 10 mL steric isotonic saline. Blood samples were drawn at intervals before, during, and after (up to 180 min) infusion to allow the coagulation state of the animal to be monitored. Plasma samples for determination of clotting times, [fibrinogen], [factor V], and [factor VIII], were anticoagulated with 0.38% citrate. Samples for measurement of aPC levels were anticoagulated with 0.38% citrate plus 25 mM *p*-aminobenzamidine. Samples before ( $t = 0$ ) and after infusion ( $t = 10$  min) for determination of fibrinopeptide A and D-dimer levels were anticoagulated with an inhibitor cocktail (containing citrate, heparin, and specific protease inhibitors) provided with the assay kit (Asserachrom FPA). Before ( $t = 0$ ) and after ( $t = 180$  min) samples for determination of platelet factor 4 and  $\beta$ -thromboglobulin were anticoagulated with 0.38% citrate plus adenosine, theophylline, and dipyridamole. Plasma was obtained from whole blood immediately. Whole blood samples for determination of platelet counts (before and after) were collected in 0.15% EDTA.

All clotting assays and quantitative determinations performed on plasma samples have been described previously (Gibbs et al., 1995). Template bleeding times were performed before infusion ( $t = 0$ ) and at the peak of anticoagulation ( $t = 20$  min). A blood pressure cuff was inflated to 40 mmHg, and a defined incision was made on the volar surface of the forearm using a commercial device (Surgicutt Jr.). The wound was blotted at 15 s intervals, and the time to the cessation of bleeding was recorded.

## RESULTS

**Initial Screening of 76 Saturation Mutants of Residues W50, K52, E229, and R233.** Each mutant was transiently expressed in COS-7 cells and assayed for TM-dependent PC

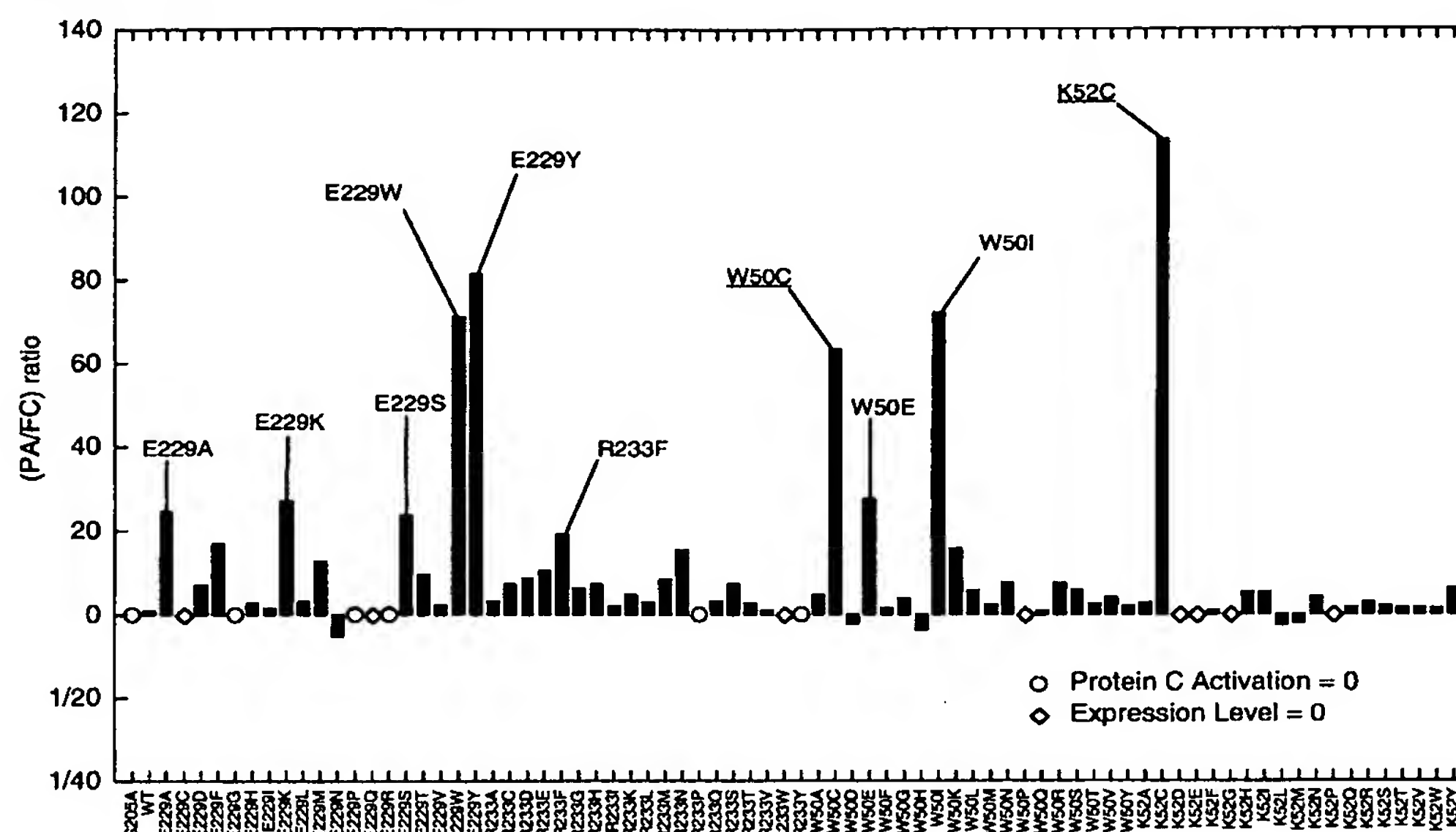


FIGURE 1: Initial screening of thrombin mutants for enhanced specificity toward protein C following saturation mutagenesis of residues W50, K52, E229, R233. PC activation (PA) and fibrinogen clotting (FC) specific activity was determined for each thrombin mutant in concentrated conditioned cell culture medium following transient transfection of COS-7 cells. Specificity was expressed as the PA/FC ratio where wild-type thrombin has a PA/FC ratio of 1. A ratio of greater than 1 indicates a greater specificity toward PC as compared to wild type (bars pointing upward), and a ratio of less than 1 indicates a lower specificity toward PC (bars pointing downward). The mean relative standard error for the PA/FC ratio of all mutants is 40%.

Table 1: Secondary Screening of Thrombin Mutants with Enhanced Protein C Activation Activity Compared to Fibrinogen Clotting Activity<sup>a</sup>

thrombin mutants <sup>b</sup>	FPA release, EC <sub>50</sub> (nM)	TM binding, K <sub>d</sub> (nM)	PC activation, <sup>c</sup> V <sub>s</sub> (pmol of aPC/h)	ATIII inhibition, IC <sub>50</sub> (μM)
WT	ND	3.46 ± 0.41	6.12 ± 0.59	0.069 ± 0.003
E229A	0.69 ± 0.05	3.75 ± 1.13	2.37 ± 0.66	1.97 ± 0.23
E229K	2.90 ± 0.16	4.38 ± 0.67	1.38 ± 0.16	5.73 ± 0.41
E229S	0.75 ± 0.02	4.69 ± 0.70	2.54 ± 0.30	1.34 ± 0.09
E229W	3.55 ± 0.30	3.43 ± 0.40	1.36 ± 0.11	0.78 ± 0.03
E229Y	0.34 ± 0.01	4.12 ± 0.66	1.80 ± 0.22	1.08 ± 0.07
R233F	0.49 ± 0.01	3.52 ± 0.46	0.74 ± 0.07	1.48 ± 0.11
W50E	0.35 ± 0.02	1.88 ± 0.47	0.31 ± 0.05	5.28 ± 0.10
W50I	0.125 ± 0.016	3.32 ± 0.44	1.36 ± 0.13	1.53 ± 0.06

<sup>a</sup> Mutants that displayed a PA/FC ratio > 18 following initial screening (Figure 1) were purified and characterized in more detail. <sup>b</sup> Errors represent standard errors. <sup>c</sup> V<sub>s</sub> = rate of PC activation for 0.5 nM TM saturated with thrombin and 1 μM PC.

activation and fibrinogen clotting activity. Specificity for protein C was expressed as the ratio of the specific activity of PC activation (PA) over the specific activity of fibrinogen clotting (FC) (Figure 1). The ratio for wild-type thrombin was assigned a value of 1. Nine mutants with a PA/FC ratio of greater than 18, E229K, E229S, E229W, E229Y, R233F, W50C, W50E, W50I, and K52C, were identified in addition to E229A thrombin, the prototype protein C activator. Because the PA/FC ratios for mutants with very low specific fibrinogen clotting activity (FC) values may be subject to error, all these mutants except W50C and K52C were selected for secondary screening. W50C and K52C were not selected for further analysis despite their favorable PA/FC ratio because of their free cysteine, which has the potential to cause dimer formation or disruption of normal disulfide bridges within the thrombin molecule.

**Secondary Screening of 7 Mutants with an Enhanced PA/FC Ratio.** The seven thrombin mutants with a PA/FC ratio > 18 identified in the initial screening were subcloned into the Sindbis virus expression vector for medium scale expression. About 100–120 μg of each mutant thrombin was purified for kinetic characterization. For secondary screening, the EC<sub>50</sub> for FPA release, the K<sub>d</sub> for TM binding,

the V<sub>s</sub> for PC activation, and the IC<sub>50</sub> of ATIII inhibition were determined (Table 1). The affinity for TM of all the mutants tested did not change significantly from that of the wild type. E229K and E229W thrombins were distinguished by their decreased activity toward fibrinogen, displaying a 4–5-fold increase in the EC<sub>50</sub> for FPA release compared to E229A thrombin. However, their V<sub>s</sub> for TM-dependent PC activation remained comparable to that of E229A thrombin. In addition, E229K but not E229W thrombin displayed increased resistance to ATIII. These two mutants were subjected to more detailed kinetic characterization to determine which had superior specificity for PC.

**Final Discrimination between E229K and E229W Thrombins.** The kinetic parameters of FPA release and PC activation were determined for wild-type, E229A, E229K, and E229W thrombins (Table 2). The k<sub>cat</sub>/K<sub>m</sub> of FPA release for E229K and E229W were almost identical and ~5-fold lower than that of E229A and ~300-fold lower than that of wild-type thrombin. The k<sub>cat</sub>/K<sub>m</sub> of PC activation in the presence of TM for E229K did not change from that of E229A while that of E229W decreased by 50.5% with respect to E229A. The net result is that E229K had a 130.9-fold higher specificity for PC over fibrinogen than wild type,



Table 2: Kinetic Analysis of Purified Thrombin Mutants with Enhanced Specificity for Protein C over Fibrinogen

thrombin mutants <sup>a</sup>	fibrinopeptide A release <sup>b</sup>			protein C activation <sup>c</sup>			specificity ratio <sup>d</sup>
	$k_{cat}$ (s <sup>-1</sup> )	$K_m$ (μM)	$k_{cat}/K_m$ (μM <sup>-1</sup> ·s <sup>-1</sup> )	$k_{cat}$ (s <sup>-1</sup> )	$K_m$ (μM)	$k_{cat}/K_m$ (μM <sup>-1</sup> ·s <sup>-1</sup> )	CE <sub>PA</sub> /CE <sub>FR</sub>
WT	59 ± 13	3.16 ± 0.38	18.7 ± 6.7	0.650 ± 0.056	3.19 ± 0.51	0.20 ± 0.05	0.0107 (1)
E229A	7.88 ± 0.03	23.6 ± 0.1	0.334 ± 0.003	0.189 ± 0.005	2.00 ± 0.13	0.095 ± 0.009	0.2844 (26.59)
E229K	2.53 ± 0.23	38 ± 3	0.07 ± 0.01	0.21 ± 0.01	2.15 ± 0.26	0.098 ± 0.016	1.4 (130.90)
E229W	3.12 ± 0.02	52.8 ± 0.3	0.059 ± 0.001	0.148 ± 0.004	3.12 ± 0.19	0.047 ± 0.004	0.7966 (74.48)

<sup>a</sup> Errors are standard errors. <sup>b</sup> The thrombin concentrations used for WT, E229A, E229K, and E229W were 0.3, 7.0, 30, and 25 nM, respectively. <sup>c</sup> Thrombin concentration was 30 nM, TM was 0.5 nM, and PC was varied from 0.5 to 10 μM. For calculation of  $k_{cat}$  the enzyme concentration was assumed to be equal to that of TM (0.5 nM) saturated with thrombin. <sup>d</sup> The specificity ratio is the ratio of the catalytic efficiency for protein C activation (CE<sub>PA</sub>) divided by the catalytic efficiency for fibrinopeptide A release (CE<sub>FR</sub>). The relative specificity compared to that of wild-type thrombin (specificity ratio<sub>mutant</sub>/specificity ratio<sub>wild type</sub>) is in parentheses.

Table 3: Further Characterization of Thrombin Mutant E229K

thrombin mutants <sup>a</sup>	platelet aggregation, EC <sub>50</sub> (nM)	S-2238 hydrolysis, $K_m$ (μM)	ATIII/heparin inhibition, $k_2 \times 10^{-8}$ (M <sup>-1</sup> ·min <sup>-1</sup> )	ATIII inhibition, $k_2 \times 10^{-5}$ (M <sup>-1</sup> ·min <sup>-1</sup> )	plasma inactivation, $t_{1/2}$ (s)
WT	1.3 ± 0.9	2.95 ± 0.46	5.98 ± 0.14	3.25 ± 0.18	36 ± 9
E229A	15 ± 7	50.2 ± 2.7	0.63 ± 0.01	0.47 ± 0.04	169 ± 48
E229K	23 ± 9	63.2 ± 4.9	0.181 ± 0.005	0.150 ± 0.004	934 ± 224

<sup>a</sup> Errors represent standard errors.

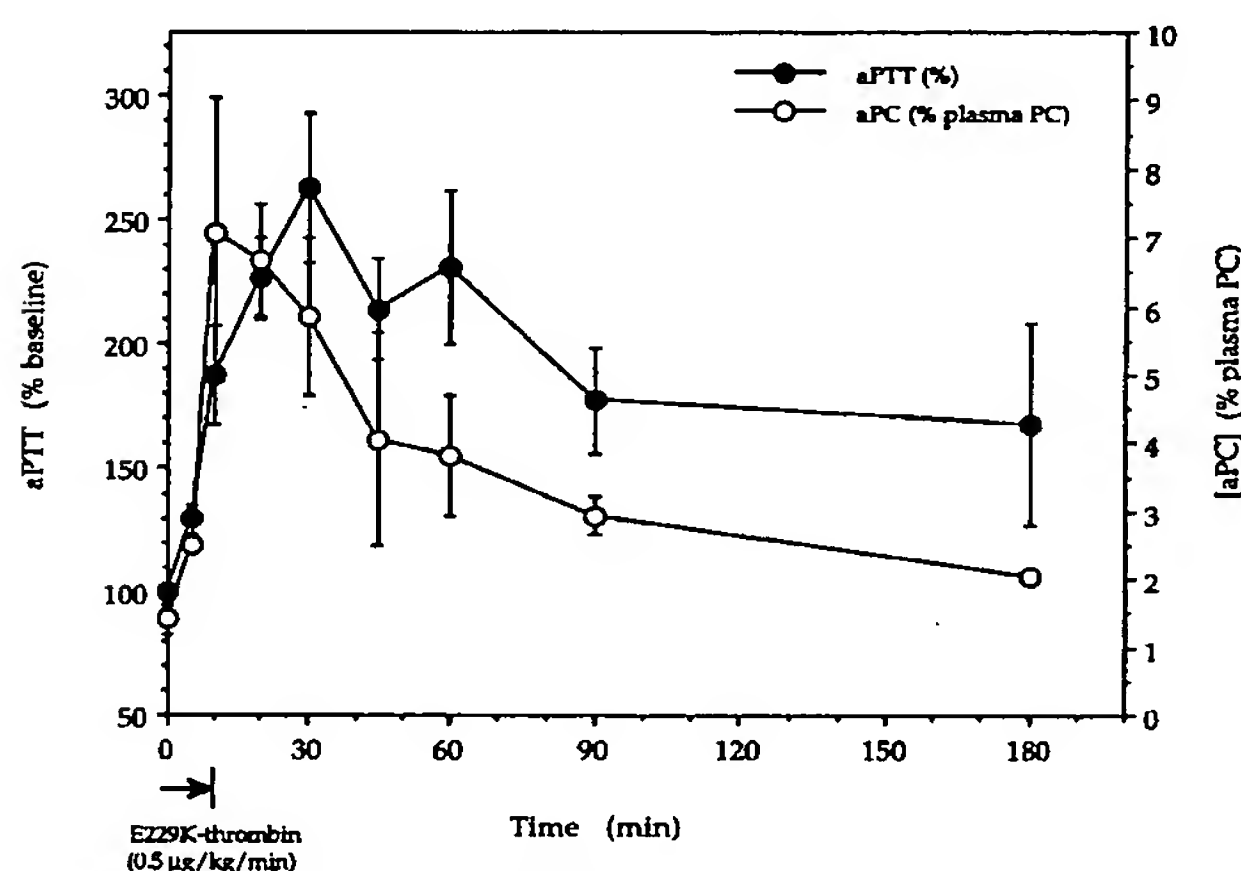


FIGURE 2: Activation of endogenous protein C and anticoagulation following infusion of E229K thrombin in Cynomolgus monkeys: E229K thrombin was infused at a dose of 0.5 μg kg<sup>-1</sup> min<sup>-1</sup> for 10 min ( $n = 4$ ). Anticoagulation was monitored by prolongation of the aPTT (baseline = 19.8 s); endogenous aPC levels were recorded as a percent of total monkey plasma PC. Error bars represent standard errors.

whereas E229W had only a 74.48-fold higher specificity for PC. E229K thrombin was also ~18-fold less efficient than wild type in its ability to activate platelets (Table 3) and also displayed increased resistance to antithrombin III compared to wild type, in both the presence and absence of heparin, by a factor of 33 and 22, respectively. The increased resistance to antithrombin III may be correlated to the observed 26-fold increase in plasma half-life *in vitro* (Table 3).

**Anticoagulation *in Vivo* following Infusion of E229K Thrombin.** Infusion of E229K thrombin in four Cynomolgus monkeys at a dose of 0.5 μg kg<sup>-1</sup> min<sup>-1</sup> for 10 min caused activation of endogenous protein C and an anticoagulant effect as measured by prolongation of the aPTT (Figure 2). The peak anticoagulant effect was observed between  $t = 10$  min and  $t = 30$  min over which period the aPTT was prolonged by approximately 2.3-fold and approximately 6.6% of total plasma PC was observed to be in the aPC form. The effect of the infusion was reversible with aPC levels returning to baseline with a half-life of approximately 60 min.

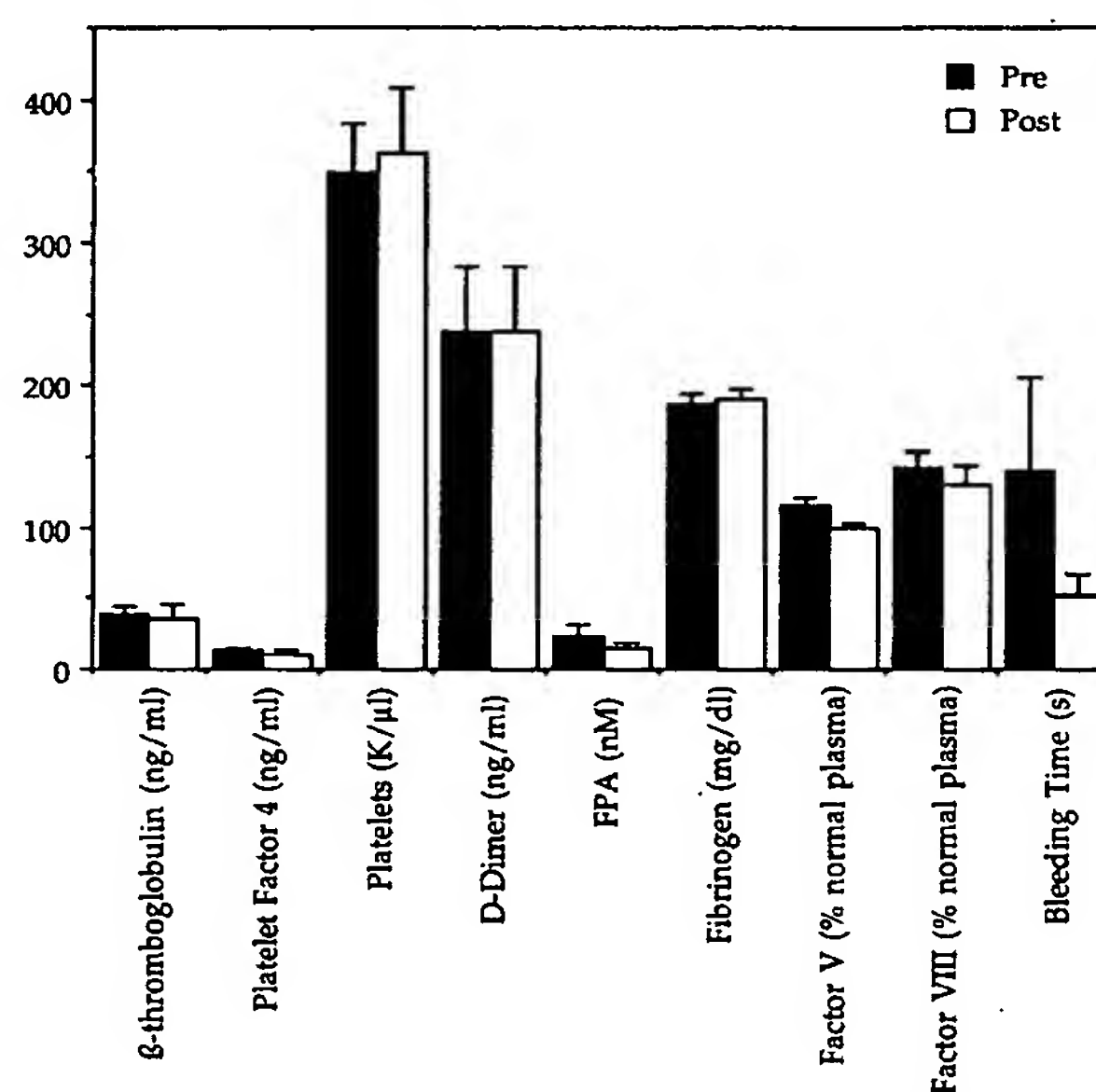


FIGURE 3: Maintenance of hemostatic parameters following infusion of E229K thrombin in Cynomolgus monkeys: Preinfusion (Pre,  $t = 0$ ) and postinfusion levels (Post) of hemostatic parameters were determined following 10 min infusions of E229K thrombin at a dose of 0.5 μg kg<sup>-1</sup> min<sup>-1</sup> ( $n = 4$ ). Postinfusion values were determined at  $t = 180$  min except for D-dimer and FPA ( $t = 10$  min) and bleeding time ( $t = 20$  min). Error bars represent standard errors.

Notably, template bleeding times measured at the peak of anticoagulation were not prolonged (Figure 3).

Fibrinogen was not consumed and FPA and D-dimer levels were not elevated following infusion of E229K thrombin (Figure 3), indicating that E229K thrombin did not cause detectable fibrinogen clotting *in vivo*. Similarly, the levels of coagulation factors V and VIII were not perturbed (Figure 3), indicating that E229K thrombin did not activate the coagulation cascade. In addition, platelets were not consumed and plasma levels of platelet secretion granule components, β-thromboglobulin and platelet factor 4, were not elevated (Figure 3), indicating that E229K thrombin did not stimulate platelet activation *in vivo*.



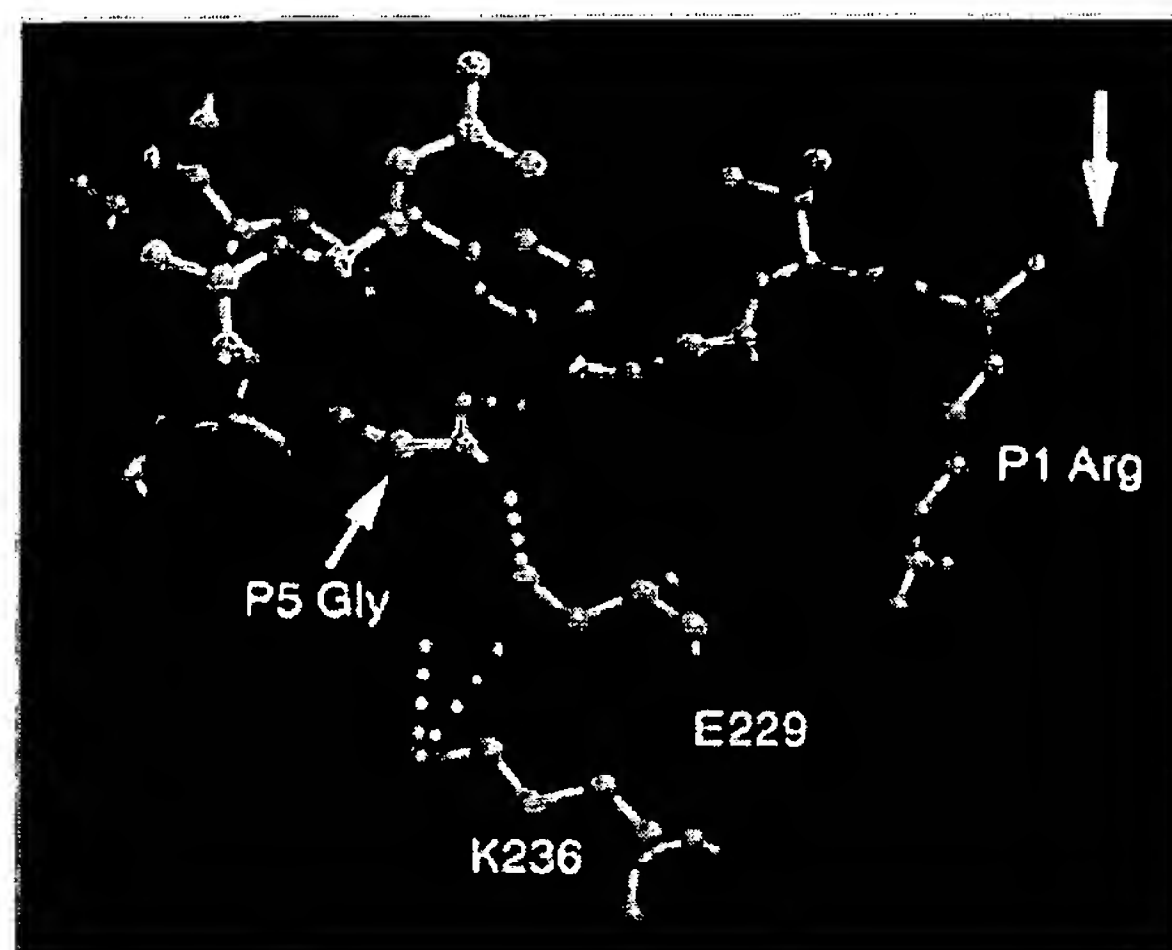


FIGURE 4: Specific contacts of the side chain of thrombin residue E229 in the complex of thrombin with FPA. The image was constructed using MidasPlus (UCSF) on the basis of the crystal structure of human thrombin complexed with FPA (Stubbs et al., 1992). Fibrinopeptide A is shown with the P1 and P5 residues highlighted. The carbonyl oxygen of the amide bond in FPA cleaved by thrombin is indicated with an arrow. The only thrombin residues shown are E229 and K236. E229 is in van der Waals contact with the glycine residue at P5 in FPA (E229 C $\gamma$ -P5 Gly O distance = 3.43 Å) and forms a salt bridge with K236 (E229 O $\epsilon$ 1-K236 N $\zeta$  = 4.11 Å; E229 O $\epsilon$ 2-K236 N $\zeta$  = 3.33 Å). Atom coloring is as follows: carbon = brown, oxygen = red, and nitrogen = blue. Interatomic contacts with the side chain of E229 are indicated by white dotted lines.

## DISCUSSION

In previous *in vivo* studies where thrombin infusion has been utilized to achieve an anticoagulant effect through the activation of endogenous PC, wild-type thrombin displayed a weak anticoagulant activity *in vivo* due to rapid clearance by ATIII and the dose-limiting effects of the familiar procoagulant functions of wild-type thrombin (fibrinogen clotting and platelet activation) (Hanson et al., 1993; Gibbs et al., 1995). In this study, screening of a library of 76 thrombin mutants, where all possible substitutions at positions W50, K52, E229, and R233 were sampled, identified a single amino acid substitution, E229K, that conferred optimal anticoagulant activity. Substitution of thrombin residue E229 with lysine resulted in a 267-fold decrease in the catalytic efficiency of cleavage of the fibrinogen A $\alpha$  chain, an 18-fold decrease in the efficiency of cleavage of the platelet thrombin receptor, and a 33- and 22-fold decrease in the rate of inhibition by ATIII in the presence and absence of heparin. In contrast, the affinity for TM was unchanged and the catalytic efficiency of TM-dependent PC activation was decreased by only 51%.

Thrombin residue E229 is located at surface of thrombin on the edge of the active site cleft where it is well positioned to interact with substrates and active site inhibitors such as ATIII on the N-terminal side of the scissile bond. In the crystal structure of human thrombin complexed with FPA, E229 is observed in van der Waals contact with a glycine residue at P5 in FPA (Figure 4). As discussed previously (Gibbs et al., 1995), the most attractive structural rationalization for the differential effects of the substitution of E229 on substrate recognition is the hypothesis that PC binds thrombin in a mode that is distinct from the other substrates and less affected by the replacement of E229. This

hypothesis gains some support from the crystal structure of human thrombin complexed with FPA, where FPA is bound in a hooked conformation (Figure 4) due to a type II  $\beta$ -turn that requires the residue at P5 to be glycine to avoid strain (Stubbs et al., 1992). In PC, the absence of glycine at P5 or elsewhere proximal to the cleavage site suggests PC is likely to bind thrombin in an extended conformation unlike that observed with FPA.

On average, the catalytic deficiencies toward the procoagulant substrates and inhibition by ATIII are approximately 5-fold more severe for the E229K variant than for E229A. This may be due to the introduction of a residue of opposite charge or altered conformation at position 229, leading to less favorable interactions with these substrates. Alternatively, the effects of the lysine substitution may be more indirect. In thrombin, the side chain of E229 can form an ion pair with the side chain of K236 (distance between charged centers = 3.7 Å) (Figure 4) (Bode et al., 1992). The introduction of another lysine residue at position 229 may cause charge repulsion and more generalized structural changes that prevent optimal interactions with the procoagulant substrates.

The effects of the improved specificity and resistance to clearance of E229K thrombin observed *in vitro* were also manifested *in vivo* upon infusion of E229K thrombin in Cynomolgus monkeys. E229K thrombin was a more potent PC activator and anticoagulant than E229A thrombin in terms of both magnitude and duration of the effects. Previously, following infusion of E229A thrombin at a dose of 2.5  $\mu\text{g kg}^{-1} \text{ min}^{-1}$  the maximal prolongation of the aPTT was 2.3-fold; aPC levels peaked at 5.5% of total plasma PC and returned to baseline with a half-life of approximately 15 min (Gibbs et al., 1995). Infusion of E229K thrombin caused effects of a similar magnitude (2.6-fold maximal prolongation of aPTT and peak conversion of 7.1% of plasma PC to aPC) but at a 5-fold lower dose of 0.5  $\mu\text{g kg}^{-1} \text{ min}^{-1}$ . In addition, aPC levels returned to baseline more slowly with a half-life of approximately 60 min, which results in a cumulative increase in potency of greater than 10-fold when the area under the decay curve is taken into account. Although it is difficult to extrapolate *in vitro* clearance studies to *in vivo* situations, the increased potency of E229K thrombin observed *in vivo* may be correlated to its prolonged half-life in plasma *in vitro* (Table 3) and its resistance to purified ATIII (Table 3), which is the most important physiological inhibitor of thrombin (Rosenberg, 1987; Olsen & Björk, 1992).

The enhanced specificity of E229K thrombin toward PC was also manifested *in vivo*. Previous studies in baboons (Hanson et al., 1993) and Cynomolgus monkeys (Gibbs et al., 1995) demonstrated that, following infusion of wild-type thrombin, the effects of both the procoagulant and the anticoagulant functions of thrombin were evident. In addition to activating PC and prolonging clotting times, wild-type thrombin also cleaved fibrinogen (fibrinogen consumption, elevation of FPA and D-dimer), activated the coagulation cascade (consumption of factor VIII and factor V), and activated platelets (elevation of platelet factor 4 and  $\beta$ -thromboglobulin). Following infusion of the prototype PC activator, E229A thrombin, in monkeys (Gibbs et al., 1995), the anticoagulant effect was observed in the absence of almost all of the aforementioned procoagulant side effects with the only indication of any procoagulant activity being a mild elevation of FPA levels (42 nM compared with 254 nM for

wild type). In the present study there was no indication of any residual procoagulant function following infusion of E229K thrombin. FPA levels remained constant at preinfusion levels (15 nM), which may reflect the enhanced specificity of E229K thrombin for PC over fibrinogen.

A notable feature of the anticoagulation studies in Cynomolgus monkeys is that anticoagulation by E229K thrombin as well as E229A thrombin (Gibbs et al., 1995) was not associated with prolonged template bleeding times even at elevated doses where the aPTT was in excess of 200 s ( $> 10$ -fold prolonged,  $n = 1$ ) (data not shown). The mechanism of anticoagulation mediated by thrombin variants is through endogenous aPC generation which by inactivating factors Va and VIIIa inhibits generation of factor Xa and thrombin. This mechanism is similar to that shared by inhibitors acting directly on coagulation factors upstream of thrombin. In contrast to direct thrombin inhibitors [PPACK (Hanson & Harker, 1988) and hirudin (Kelly et al., 1991)] and direct antiplatelet agents [mAbs versus vWF and GPIIb/IIIa (Cadroy et al., 1994) and RGDV peptide (Cadroy et al., 1989)], anticoagulants that act by inhibiting thrombin generation [tick anticoagulant peptide (Sitko et al., 1992) and aPC (Gruber et al., 1990, 1991)] have shown antithrombotic efficacy without perturbing primary hemostasis as measured by bleeding time.

Therefore, E229K thrombin which can function as a potent anticoagulant *in vivo* without any procoagulant side effects and without prolonging bleeding time may be a superior antithrombotic agent with reduced potential for bleeding complications.

#### ACKNOWLEDGMENT

We thank Pam Mack, DVM (Tulane Regional Primate Research Center) for assistance with animal studies and Terry Terhorst for the synthesis of the oligonucleotides used in this study.

#### REFERENCES

- Bajzar, L., Manuel, R., & Nesheim, M. E. (1995) *J. Biol. Chem.* 270, 14477–14484.
- Bode, W., Turk, D., & Karshikov, A. (1992) *Protein Sci.* 1, 426–471.
- Bredenbeek, P. J., Frolov, I., Rice, C. M., & Schlesinger, S. (1993) *J. Virol.* 67, 6439–6446.
- Cadroy, Y., Houghton, R. A., & Hanson, S. R. (1989) *J. Clin. Invest.* 84, 939–944.
- Cadroy, Y., Hanson, S. R., Kelly, A. B., Marzec, U. M., Evatt, B. L., Kunicki, T. J., & Montgomery, R. R. (1994) *Blood* 83, 3218–3224.
- Comp, P. C., & Esmon, C. T. (1981) *J. Clin. Invest.* 68, 1221–1228.
- Comp, P. C., Jacocks, R. M., Ferrell, G. L., & Emson, C. T. (1982) *J. Clin. Invest.* 70, 127–134.
- Esmon, C. T. (1987) *Science* 235, 1348–1352.
- Fenton, J. W., II, Fasco, M. J., & Stackrow, A. B. (1977) *J. Biol. Chem.* 252, 3587–3598.
- Frieznier Degen, S. J., MacGillivray, R. T. A., & Davie, E. W. (1983) *Biochemistry* 22, 2087–2097.
- Gailani, D., & Broze, G. J., Jr. (1991) *Science* 253, 909–912.
- Gibbs, C. S., & Zoller, M. J. (1991) *Methods: Companion Methods Enzymol.* 3, 165–173.
- Gibbs, C. S., Coutre, S. E., Tsiang, M., Li, W.-X., Jain, A. K., Dunn, K. E., Law, V. S., Mao, C. T., Matsumura, S. Y., Mejza, S. J., Paborsky, L. R., & Leung, L. L. K. (1995) *Nature* 378, 413–416.
- Gomi, K., Zushi, M., Honda, G., Kawahara, S., Matsuzaki, O., Kanabayashi, T., Yamamoto, S., Maruyama, I., & Suzuki, K. (1990) *Blood* 75, 1396–1399.
- Graycar, T. P., & Estell, D. A. (1987) *J. Cell. Biochem., Suppl.* 11c, 234.
- Gruber, A., Griffin, J. H., Harker, L. A., & Hanson, S. R. (1989) *Blood* 73, 639–642.
- Gruber, A., Hanson, S. R., Kelly, A. B., Yan, B. S., Bang, N., Griffin, J. H., & Harker, L. A. (1990) *Circulation* 82, 578–585.
- Hanson, S. R., & Harker, L. A. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85, 3184–3188.
- Hanson, S. R., Griffin, J. H., Harker, L. A., Kelly, A. B., Emson, C. T., & Gruber, A. (1993) *J. Clin. Invest.* 92, 2003–2012.
- Higgins, D. L. (1983) *J. Biol. Chem.* 258, 9276–9282.
- Jackson, C. M., & Nemerson, Y. (1980) *Annu. Rev. Biochem.* 49, 765–811.
- Kelly, A. B., U. M. M., Krupski, W., Bass, A., Cadroy, Y., Hanson, S. R., & Harker, L. A. (1991) *Proc. Natl. Acad. Sci. U.S.A.* 77, 1006–1012.
- Kogan, A. E., Bashkov, G. V., Bobruskin, I. D., Romanova, E. P., Makarov, V. A., & Strukova, S. M. (1993) *Thromb. Res.* 70, 385–393.
- Kurz, K. D., Smith, T., Wilson, A., Gerlitz, B., Richardson, M. A., & Grinnell, B. W. (1994) *Circulation* 90, 1–180.
- Liljeström, P., & Garoff, H. (1991) *Bio/Technology* 9, 1356–1361.
- Liljeström, P., Lusa, S., Huylebroeck, D., & Garoff, H. (1991) *J. Virol.* 65, 4107–4113.
- Lorand, L., & Radek, J. T. (1992) In *Thrombin Structure and Function* (Berliner, L. J., Ed.) pp 257–270, Plenum Press, Columbus, OH.
- Mann, K. G., & Lundblad, R. L. (1987) In *Hemostasis and Thrombosis* (Colman, R. W., Hirsh, J., Marder, V. J., & Salzman, E. W., Eds.) pp 148–161, Lippincott, Philadelphia, PA.
- Mann, K. G., Jenny, R. J., & Krishnaswamy, S. (1988) *Annu. Rev. Biochem.* 57, 915–956.
- McBane, R. D., Wysokinski, W. E., & Chesebro, J. H. (1995) *Thromb. Haemostasis* 74, 879–885.
- Okajima, K., Imamura, H., Koga, S., Inoue, M., Takatsuki, K., & Aoki, N. (1990) *Am. J. Hematol.* 33, 277–278.
- Olson, S. T., & Bjork, I. (1992) In *Thrombin Structure and Function* (Berliner, L. J., Ed.) pp 159–217, Plenum Press, New York.
- Rice, C. M., Levis, R., Strauss, J. H., & Huang, H. V. (1987) *J. Virol.* 61, 3809–3819.
- Richardson, M. A., Gerlitz, B., & Grinnell, B. W. (1992) *Nature* 360, 261–264.
- Rosenberg, R. D. (1987) In *The Molecular Basis of Blood Diseases* (Stamatoyannopoulos, G., Nienhuis, A. W., Leder, P., & Majerus, P. W., Eds.) pp 534–574, W. B. Saunders Co., Philadelphia, PA.
- Sheehan, J. P., Wu, Q., Tollefsen, D. M., & Sadler, J. E. (1993) *J. Biol. Chem.* 268, 3639–3645.
- Sitko, G. R., Ramjit, D. R., Stabilito, I. L., Lehman, D., Lynch, J. J., & Vlasuk, G. P. (1992) *Science* 258, 593–596.
- Strukova, S. M., Kogan, A. E., Tara, A., & Aaviksaar, A. (1989) *Thromb. Res.* 55, 149–153.
- Stubbs, M. T., Oschkinat, H., Mayr, I., Huber, R., Angliker, H., Stone, S. R., & Bode, W. (1992) *Eur. J. Biochem.* 206, 187–195.
- Taylor, F. B., Jr., Chang, A., Esmon, C. T., D'Angelo, A., Vigano-D'Angelo, S., & Blick, K. E. (1987) *J. Clin. Invest.* 79, 918–925.
- Tsiang, M., Lentz, S. R., Dittman, W. A., Wen, D., Scarpatti, E. M., & Sadler, J. E. (1990) *Biochemistry* 29, 10603–10612.
- Tsiang, M., Lentz, S. R., & Sadler, J. E. (1992) *J. Biol. Chem.* 267, 6164–6170.
- Tsiang, M., Jain, A. K., Dunn, K. E., Rojas, M. E., Leung, L. L. K., & Gibbs, C. S. (1995) *J. Biol. Chem.* 270, 16854–16863.
- van Hinsbergh, V. W. M., Bertina, R. M., van Wijngaarden, A., van Tilburg, N. H., Emeis, J. J., & Haverkate, F. (1985) *Blood* 65, 444–451.
- Vu, T.-K. H., Hung, D. T., Wheaton, V. I., & Coughlin, S. R. (1991) *Cell* 64, 1057–1068.
- Wu, Q., Sheehan, J. P., Tsiang, M., Lentz, S. R., Birktoft, J. J., & Sadler, J. E. (1991) *Proc. Natl. Acad. Sci. U.S.A.* 88, 6775–6779.
- Ye, J., Rezaie, A. R., & Esmon, C. T. (1994) *J. Biol. Chem.* 269, 17965–17970.



# Affinity Maturation of Human Growth Hormone by Monovalent Phage Display

Henry B. Lowman and James A. Wells†

Department of Protein Engineering  
Genentech, Inc., 460 Pt San Bruno Boulevard  
South San Francisco, CA 94080, U.S.A.

(Received 3 June 1993; accepted 30 June 1993)

We describe a selection procedure for construction of very high affinity variants of human growth hormone (hGH) for binding to the extra cellular domain of its receptor (called the hGHbp). Five different libraries of mutated hGH genes (each containing  $\sim 2 \times 10^5$  protein variants) were created by randomly mutating four different codons at residues that were shown to be important for receptor binding by structural or functional criteria. Mutated proteins were displayed as single copies from their respective filamentous phagemid particles and sorted *in vitro* for binding to the immobilized hGHbp. Phagemid particles that bound the immobilized hGHbp were eluted and propagated. After three to seven rounds of binding enrichments, hGH variants were isolated that contained 2 to 4 mutations and exhibited three- to sixfold improvements in binding affinity.

Because of the limits of DNA transfection efficiency in creating the library we could not sample thoroughly mutations at more than four codons at once. Nonetheless, the free energy effects for these mutations acted cumulatively. Thus, by combining affinity enhanced mutants from libraries independently sorted we created an hGH variant with 15 substitutions that bound  $\sim 400$ -fold more tightly to the hGHbp than wild-type hGH. The affinity enhancements occurred predominantly by slowing the off-rate of the hormone ( $> 60$ -fold), and partly through increasing the on-rate (up to 4-fold). Residues that were shown to be important for binding by alanine-scanning were most highly conserved after binding selection, and interestingly many of them could be further improved. Thus, we found it most effective to randomly mutate the residues that were shown to modulate affinity by alanine-scanning, and to combine the selectants from separate libraries that exhibit the highest affinities. The selection procedure and mutagenesis strategy provides a framework for affinity maturation of protein-protein complexes.

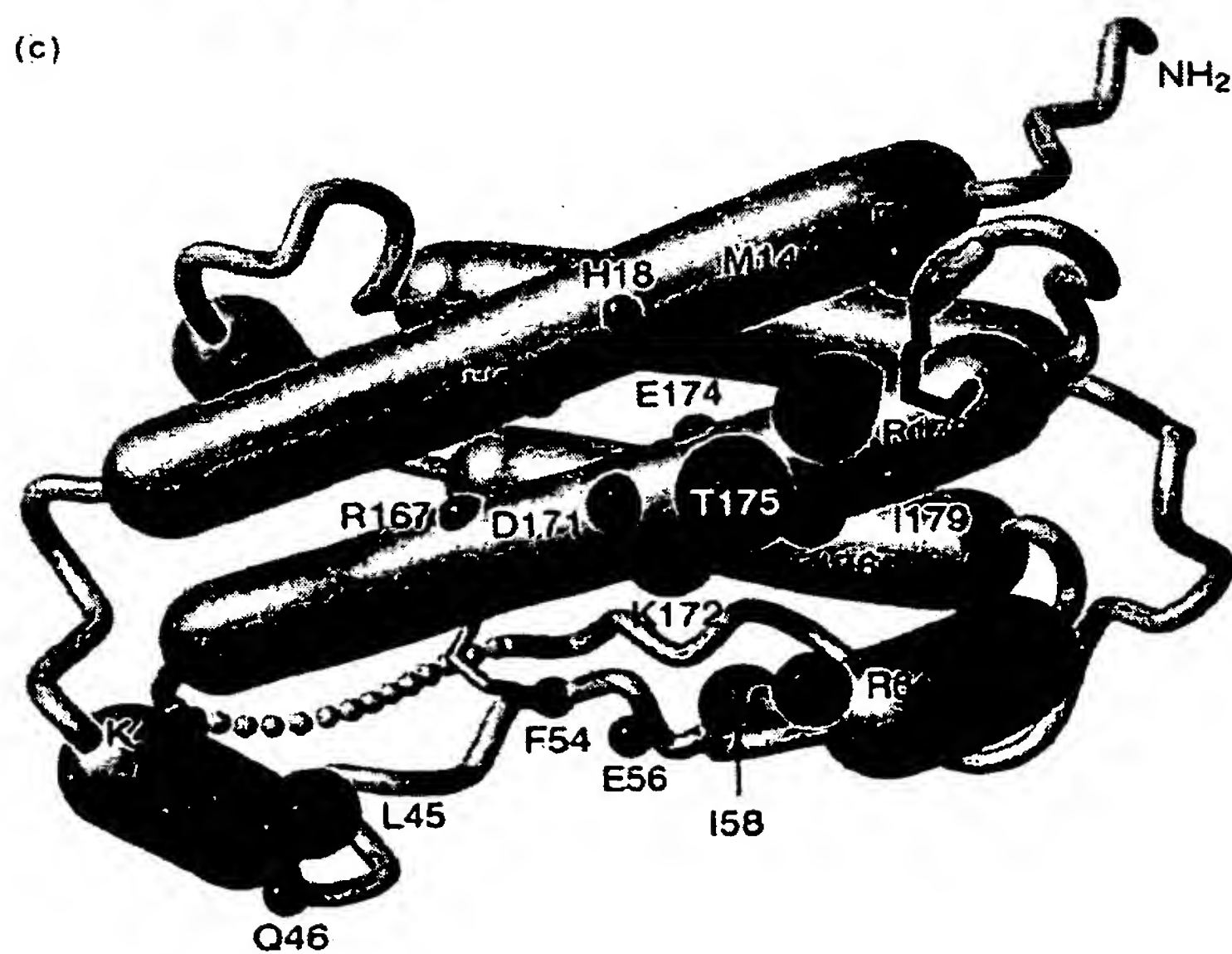
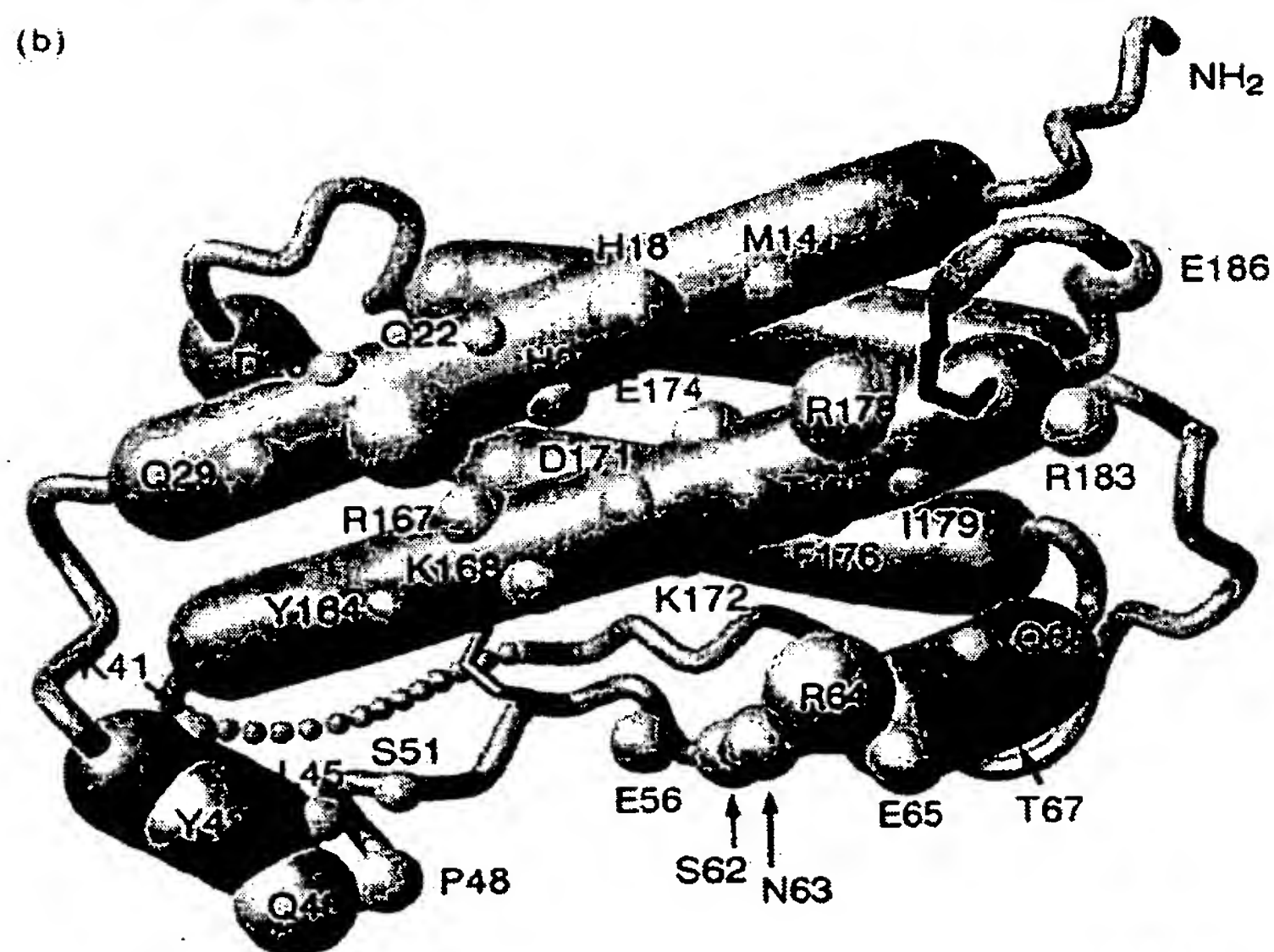
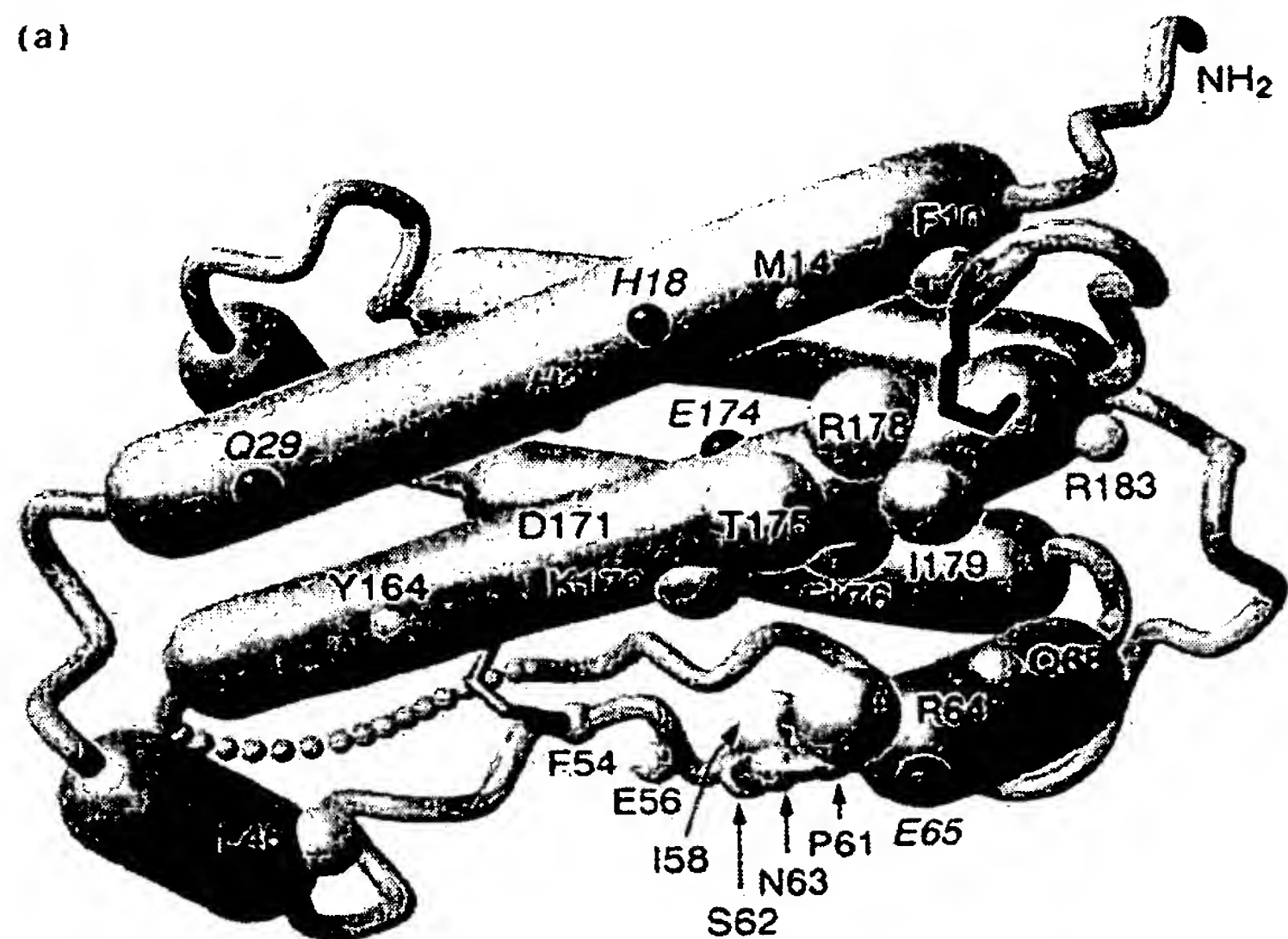
**Keywords:** mutagenesis; phage display; human growth hormone; structure/function; hormones/receptors

† Author to whom all correspondence should be addressed.

‡ Abbreviations used: hGH, human growth hormone; hGHbp, extracellular domain of the hGH receptor also called the hGH binding protein; hPL, human placental lactogen; EDC, *N*-ethyl-*N'*-(dimethylaminopropyl)-carbodiimide; NHS, *N*-hydroxysuccinimide; PBS, phosphate buffered saline (pH 7.4); PDEA, 2-(2-pyridinyldithio)ethaneamine hydrochloride; MAb, monoclonal antibody; RIA, radio-immunoprecipitation assay; ssDNA and dsDNA, single-stranded and double-stranded DNA; SPR, surface plasmon resonance. hGH mutants are designated by the wild-type residue (in single-letter code) followed by its position in the sequence and the mutant residue. For example, R64K indicates a mutation in which Arg64 is converted to Lys. Multiple mutants are indicated by a series of single mutants separated by slashes.

## 1. Introduction

Determining how proteins associate is fundamental to molecular recognition processes in biology. One of the most well-characterized protein-protein complexes is that between hGH† and the extracellular domain of its receptor, called the hGHbp (for review, see Wells & de Vos, 1993). High resolution structural and mutational analyses show that one hGH binds two hGHbps sequentially using two distinct sites on the hormone, called sites 1 and 2 (Cunningham *et al.*, 1991a; de Vos *et al.*, 1992). A striking feature of this complex is that less than half of the side-chains from hGH that are buried in site 1 affect binding affinity when converted to alanine (Fig. 1(a), (b); Cunningham & Wells, 1993). In fact, a few of the buried side-chains even





enhance binding affinity when converted to alanine (Cunningham & Wells, 1989, 1993; Fuh *et al.*, 1992). An hGH mutant enhanced in affinity for site 1 and blocked in its ability to bind site 2 was a better antagonist of the hGH receptor than the site 2 mutant alone (Fuh *et al.*, 1992). Thus, by improving site 1 affinity further, it may be possible to obtain an even better antagonist that could have utility in treating conditions of growth hormone excess such as acromegaly.

Monovalent phage display of hGH (Bass *et al.*, 1990) has been used to produce three to eightfold increases in the affinity of site 1 for the hGHbp (Lowman *et al.*, 1991a). Additional improvements might be obtained by mutating more residues in the phage display library. However, it was not feasible to generate enough transformants to ensure that all possible residue combinations were represented when more than five or six codons were randomized at once (see Lowman & Wells, 1991). Mutations at protein-protein interfaces usually exhibit additive effects upon binding (Wells, 1990). Therefore, we reasoned that much larger improvements in affinity may be possible by combining affinity enhanced variants from libraries in which other residues at the interface were mutated independently.

Here, very high affinity variants of hGH were produced by combining mutants of hGH that were independently selected for increased binding affinity by monovalent phage display. By this approach we produced an hGH variant containing 15 substitutions that bound the hGHbp ~400-fold tighter than wild-type hGH. These studies suggest guidelines for targeting residues for affinity optimization.

## 2. Materials and Methods

### (a) General procedures

Restriction enzymes, polynucleotide kinase, bacteriophage T<sub>7</sub> DNA polymerase, and T<sub>4</sub> DNA ligase were obtained from Gibco-BRL or New England Biolabs and used according to the manufacturer's directions. Oligonucleotide cassettes that contained random substitutions were phosphorylated, annealed, and ligated into constructs as described (Lowman *et al.*, 1991a; Lowman & Wells, 1991). Sequenase<sup>®</sup> was purchased from United States Biochemical and used according to the manufacturer's directions for single-stranded sequencing (Sanger *et al.*, 1977).

Site-specific mutants of hGH were constructed by oligonucleotide-directed mutagenesis, using a single-stranded

template that contained deoxyuracil (Kunkel *et al.*, 1991). The plasmid, pHGHam-g3 which encodes wild-type hGH fused to the carboxy-terminal domain of M13 gene III (Lowman *et al.*, 1991a), was used to construct parental vectors for cassette mutagenesis (Wells *et al.*, 1985). Phagemid particles that display a single copy of the hGH-gene III fusion protein were prepared (Lowman & Wells, 1991) by electro-transforming *Escherichia coli* XL1-Blue cells (from Stratagene<sup>®</sup>), and adding M13K07 helper phage (Vieira & Messing, 1987) which provides a large excess of the wild-type gene III protein.

Soluble hormones were expressed in *E. coli* (Chang *et al.*, 1987) and ammonium-sulfate precipitated from osmotically shocked cell supernatants (Olson *et al.*, 1981). The protein concentration was determined by laser densitometry of Coomassie-stained SDS-PAGE gels as previously described (Cunningham *et al.*, 1990). Some variants were purified further by ion-exchange chromatography on a Mono-Q column (Pharmacia-LKB Biotechnology, Inc.).

### (b) Preparation of hGH-phagemid libraries

To construct the Minihelix-1 library of hGH (M1) in which residues K41, Y42, L45 and Q46 were randomly mutated (Fig. 1(c)), the existing *Aat*II site in pHGHam-g3 was destroyed using oligonucleotide no. 718 (5'-GCCACC-TGATGTCTAAGAAAC-3'). The underlined nucleotide indicates a mismatch on the template. Unique *Sfi*I and *Aat*II sites were introduced into pHGHam-g3 to create pH0779, using oligonucleotides 782 (5'-TTTGAAGAGCC-CTATATGGCCAAGGAACAGAAG-3'), and 821 (5'-CAG-AACCCCATTTGACGTCCCTCTGTTTC-3'), respectively. The latter oligonucleotide introduced a +2 frameshift and a TGA stop codon after residue 49. A randomly mutated cassette was constructed from the complementary oligonucleotides 822 (5'-TCCCGAAGGAGCAGN-NSNNSTCGTTTNNNSNAACCCGACAGCT-3') and 823 (5'-CTCGGGGTTSSNNNSNGAACGASNNNSNCTCC-TCCCTTCGGGATAT-3'), where N represents any of the 4 nucleotides and S represents either G or C. The parental DNA (pH0779) was digested with restriction enzymes *Sfi*I and *Aat*II, and the large fragment was purified and ligated with the cassette. The ligation products were electro-transformed into XL1-Blue<sup>®</sup> cells for phagemid preparation in 2 aliquots, yielding  $1 \times 10^6$  independent transformants each as described (Lowman & Wells, 1991).

To construct the loop-1 library of hGH (L1) in which residues F54, E56, I58 and R64 were randomized, the existing *Aat*II site in pHGHam-g3 was destroyed using oligonucleotide 718. Unique *Aat*II and *Bst*EII restriction sites were introduced in the hGH gene to construct the plasmid, pH0709, using oligonucleotides 719 (5'-AACCC-CAGACGTCCCTCTGT-3') and 720 (5'-GAAACACAACA-GTAAAGGTAACCTAGAGCTGCT-3'). The latter oligonucleotide introduced a +1 frameshift and a TAA

**Figure 1.** Comparison of receptor binding epitopes defined by alanine-scanning (a), X-ray structural analysis (b), or phage display (c). (a) Residues where alanine substitutions caused a 2-fold or greater effect on receptor binding affinity on a cartoon model of hGH based on the crystal structure of the hGH(hGHbp)<sub>2</sub> complex (de Vos *et al.*, 1992). Dark spheres (●) show Ala substitutions that improved binding  $\Delta\Delta G = -1$  to  $-0.5$  kcal/mol. Light spheres (○, ○, ○, ○) denote Ala substitutions (or a Gln substitution in the case of Lys41) that reduced binding energy by  $+0.5$  to  $1.0$  kcal/mol,  $+1.0$  to  $1.5$  kcal/mol,  $+1.5$  to  $2.0$  kcal/mol, or  $+2.0$  to  $2.5$  kcal/mol, respectively. (b) The hGH site 1 structural epitope taken from Cunningham & Wells (1993). Light spheres (○, ○, ○, ○) represent a change in solvent accessible area of  $-20$  to  $0$  Å<sup>2</sup>,  $0$  to  $20$  Å<sup>2</sup>,  $20$  to  $40$  Å<sup>2</sup> or  $40$  to  $60$  Å<sup>2</sup> respectively, at each residue upon alanine substitution as calculated from the hGH(hGHbp)<sub>2</sub> crystal structure (de Vos *et al.*, 1992). (c) The conservation of residues selected for binding to the hGHbp from random phagemid libraries. The fraction of wild-type hGH residues found at each position after sorting for binding to the hGHbp is indicated by the size of the dark sphere: (●) 0 to 10% conserved; (●) 10 to 25%; (●) 25 to 50%; (●) >50% (taken from data reported herein, and Lowman *et al.*, 1991a).

stop codon after residue 69. In addition, the unique *EcoRI* site was destroyed using oligonucleotide 536 (5'-CGTCTTCAAGAGTTCAACTTCTCC-3'), to permit restriction-selection against possible contaminating clones from previous libraries (Lowman & Wells, 1991). A randomized cassette was constructed from the complementary oligonucleotides 803 (5'-pCCCTCTGTNNSTCA-NNSTCTNNSCCGACACCCAGTAATNNSGAGGAAACA-CAACAGAAGA-3') and 804 (5'-pGTTACTCTTCTGTTG-TGTTTCTCSNNATTACTGGGTGTCGGSNNAGASN-NTGASNNACAGAGGGACGT-3'). The pH0709 DNA was digested with restriction enzymes *AatII* and *BstEII*, and the large fragment was purified and ligated with the cassette. The ligation products were electrotransformed into XL1-Blue cells for phagemid preparation in 2 aliquots, yielding  $1.6 \times 10^6$  and  $1.0 \times 10^6$  independent transformants.

#### (c) Combinatorial hGH libraries

DNA from the helix-1 (called H1) library (randomly mutated at F10, M14, H18 and H21) and the helix-4B (called H4b) library (randomly mutated at R167, D171, T175, I179) in which the E174S/F176Y mutations were fixed, was isolated after binding selections for 0, 2 or 4 rounds on hGHbp-heads (Lowman *et al.*, 1991a). The DNA from each pool was purified and digested with the restriction enzymes *AclI* and *BstXI*. The large fragment from each H1 pool was purified and ligated with the small fragment from each H4b pool to yield the 3 combinatorial libraries H1.0/H4b.0 (unselected H1 and H4b pools), H1.2/H4b.2 (twice-selected H1 pool with twice-selected H4b pool), and H1.4/H4b.4 (4-times selected H1 pool with 4-times selected H4b pool).

The ligation products were processed and electrotransformed into XL1-Blue<sup>®</sup> cells. The number of independent transformants obtained from each pool was determined by colony-forming units (c.f.u.) as follows:  $2.4 \times 10^6$  for the H1.0/H4b.0 library,  $1.8 \times 10^6$  for the H1.2/H4b.2 library,  $1.6 \times 10^6$  for the H1.4/H4b.4 library. hGH-phagemid particles were prepared and selected for hGHbp-binding over 2 to 7 cycles as described (Lowman *et al.*, 1991a).

We combined the 3 highest affinity variants of hGH that were isolated after independent sorting from the H1 and H4b libraries (Lowman *et al.*, 1991a). From the H1 library these included F10H/M14G/H18N/H21N (variant H1.4A which corresponds to clone A isolated after 4 rounds of selection from the H1 library), F10A/M14W/H18D/H21N (variant H1.4B), M14S/H18F/H21L (variant H1.6B). From the H4b library these included R167N/D171S/E174S/F176Y/I179T (variant H4b.4A), R167E/D171S/E174S/F176Y (variant H4b.6E) and R167N/D171N/E174S/F176Y/I179T (variant H4b.6B). hGH-phagemid DNA was purified and digested with the restriction enzymes *EcoRI* and *BstXI*. The large fragment from each H4b variant was then purified and ligated with the small fragment from each H1 variant to yield a recombinant with mutations in both helix-1 and helix-4.

Oligonucleotides were used to revert many of the phage-derived mutations in the H1.4B/H4b.4A variant to the corresponding wild-type residue 797 (5'-CTGCCGTGC-TACCGTCTTCAACAGTTGGCCTTTG-3') for D18H/N21H; 798 (5'-GTCAGCACATTCCTGCCGACC-3') for Y176F; 799 (5'-CTCTCGCGGCTCTTCGACAACGGATG-CTGCGTGCT-3') for A10F/W14M; 800 (5'-TACTGCTTCAAGGACATGCAACGTCAGC-3') for N167R/S171D; 801 (5'-CTGCGCATCGTGCAAGTGC-3') for T179I;

875 (5'-CTCTCGAGGCTCTTCGACAACGCGTGG-3') for A10F).

The hGH variant (H1.4B/H4b.4A/M1.5B/L1.6E) was constructed using H1.4B/H4b.4A as template and the following oligonucleotides: 843 (5'-CAGACCTCCCTCTGT-CCCTCAGAGTCTATTCCG-3') for adding F54P; 844 (5'-ACACCCTCCAACAAGGAGGAAACACAACAG-3') for R64K; 846 (5'-CCAAAGGAACAGATTTCATTTCATTCTGG-TGGAACCCGAGACCTCC-3') for K41I/Y42H/L45W/Q46W. Variant M1.5B/L1.6E was constructed using the same oligonucleotides with template pHam-g3.

#### (d) Radio-immunoprecipitation assays

To determine the equilibrium binding affinity for the hGHbp, hGH variants were assayed in competition with <sup>125</sup>I-labeled hGH, or the labeled H1.4B/H4b.4A variant, or the labeled H1.4B/H4b.4A/M1.5B/L1.6E variant, in binding buffer that contained 50 mM Tris (pH 7.5) 10 mM MgCl<sub>2</sub>, 0.1% (w/v) bovine serum albumin, 0.02% (w/v) sodium azide (Lowman *et al.*, 1991b). Although these mutants contain the F176Y mutation, control experiments showed that iodination did not affect their binding affinities. Immunoprecipitation of the hGH-hGHbp complex (Cunningham *et al.*, 1991a) was carried out using MAb5 (Barnard *et al.*, 1984).

#### (e) Kinetics assays

Association and dissociation rate constants for binding of hGH variants to the hGHbp were measured by surface plasmon resonance (SPR) on a Pharmacia BIAcore<sup>™</sup> (Löfås & Johnsson, 1990). The S210C mutant of the hGHbp was immobilized on the biosensor chip *via* its free thiol group (Cunningham & Wells, 1993). This uniformly oriented the hGHbp and prevented receptor dimerization on the matrix so that site 1 binding kinetics could be tested alone. Binding and elution steps were carried out at a flow rate of 3 to 20  $\mu$ l/min in PBS buffer (pH 7.4) containing 0.05% (w/v) Tween-20.

The density of the hGHbp coupled to the matrix can affect the absolute  $k_{on}$  and  $k_{off}$  values by up to 2-fold for wild-type hGH. Thus each biochip was standardized with wild-type hGH, so that relative changes in  $k_{on}$  and  $k_{off}$  values between different mutants could be compared more accurately. Calculated affinity measurements correlated well with the results of the radio-immunoprecipitation assay (Cunningham & Wells, 1993). Dissociation rate constants were obtained by plotting  $\ln(R_0/R_t)$  versus  $t$ ; association rate constants were obtained by plotting [slope of  $(dR_t/dt)$  versus  $R_t$ ] against hormone concentration (Karlsson *et al.*, 1991), or by plotting  $\ln(dR_t/dt)$  against hormone concentration using the BIAcore<sup>™</sup> kinetics evaluation software Pharmacia Biosensor Manual). Equilibrium dissociation constants,  $K_d$ s, were calculated as  $k_{off}/k_{on}$ . Standard deviations,  $\sigma_{on}$  for  $k_{on}$  and  $\sigma_{off}$  for  $k_{off}$ , were obtained from measurements with 2 or more series of 2-fold or 3-fold dilutions ( $k_{on}$ ) or with 2 or more concentrated ( $\geq 5 \mu$ M) hormone samples ( $k_{off}$ ) (for a discussion see Bevington, 1969).

### 3. Results

#### (a) Residues in the hGH-receptor binding epitope

Structural analysis of the hGH(hGHbp)<sub>2</sub> complex (de Vos *et al.*, 1992) identified 31 side-chains in site 1 of hGH that undergo some degree of burial when



**Table 1**  
Receptor binding affinities for alanine mutants of hGH

Variant	Number of van der Waals contacts	$K_d$ (pM)	$\frac{k_d(\text{mut})}{K_d(\text{hGH})}$
hGH (wild-type)	—	340	(1)
K41A	—	NE	NE
K41Q†	7	$880 \pm 84$	2.6
Y42A	30	$540 \pm 80$	1.6
L45A	7	$3400 \pm 330$	10.0
Q46A	16	$320 \pm 20$	0.9

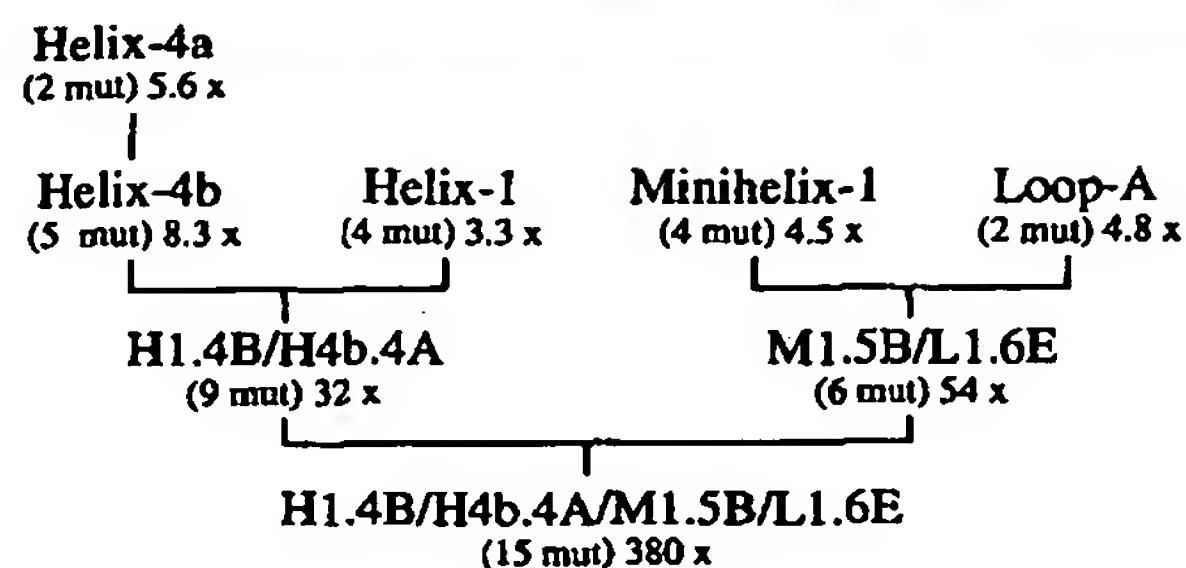
Receptor binding affinities for alanine mutants of hGH measured by BIAcore™ (†) or by RIA and normalized relative to the RIA value for wild-type hGH (Cunningham *et al.*, 1989). In the case of K41, the Ala variant did not express (NE), and so the Gln substitution was tested (Jin *et al.*, 1992). For comparison with the structural epitope, the number of van der Waals contacts with receptor is also shown, derived from the crystal structure of the hGH(hGHbp)<sub>2</sub> complex (de Vos *et al.*, 1992).

the first receptor binds (Fig. 1(b)). Although most of these were tested as alanine mutants prior to the structural elucidation (Cunningham & Wells, 1989, 1991), four residues (K41, Y42, L45 and Q46) in the first minihelix (minihelix-1) were not evaluated. We therefore converted these residues singly to alanine and measured the effects on binding affinity either by competitive displacement with [<sup>125</sup>I]hGH and immunoprecipitation (Cunningham & Wells, 1989) or using the BIAcore™ biosensor from Pharmacia. Both methods gave comparable affinity measurements (Cunningham & Wells, 1993).

The side-chains of Tyr42 and Gln46 make many van der Waals contacts with the hGHp, yet alanine replacements caused less than a twofold reduction in affinity (Table 1). The L45A mutant causes a tenfold reduction in affinity even though Leu45 makes fewer contacts with the receptor than either Tyr42 or Gln46. Lys41 makes a salt-bridge with Glu27 of the receptor. Although the K41A mutant did not express well enough to obtain material for an affinity measurement, we were able to express a more conservative variant, K41Q. This had a 2.6-fold lower affinity than wild-type hGH. With these data and those of other studies (Cunningham & Wells, 1989, 1993) the binding effects have been measured for at least one replacement (mostly alanines) for all residues whose side-chains become buried when the first receptor binds at site 1 (Fig. 1(a)).

#### (b) Design and analysis of random mutant libraries

We sorted five separate libraries (Fig. 2) in which four residues within the structural or functional site 1 epitope (Fig. 1(a), (b)) were randomly mutated. Restricting each library to four random codons allowed a reasonable sampling of the  $\sim 2 \times 10^5$  protein sequences generated from  $\sim 1 \times 10^6$  DNA sequences within an average library that contained  $\sim 5 \times 10^6$  independent transformants.



**Figure 2.** Strategy for combining mutations selected from various libraries by phage display that increase receptor binding affinity. Each library is indicated followed by the number of substitutions for the highest affinity variant isolated (in parentheses) and the increase in binding affinity over wild-type hGH. Libraries were produced by randomly mutating 4 residues at a time as follows: helix-1 [F10, M14, H18, H21]; minihelix-1 [K41, Y42, L45, Q46]; loop-1 [F54, E56, I58, R64]; helix-4a [K172, E174, F176, R178]; helix-4b [R167, D171, T175, I179]. These were sorted for binding to the hGHbp and the tightest binding variant was combined with those from other libraries to produce even higher affinity variants as described in the text and Table 4.

Previously, a library (called helix-4a; H4a) was produced in which residues K172, E174, F176 and R178 were randomized and displayed on monovalent phagemid particles (Lowman *et al.*, 1991a). After only one cycle of binding selection, a tighter binding mutant (E174S/F176Y, called H4a.1G) was isolated with an affinity about fivefold higher than wild-type hGH. The double mutant was fixed in a second library (called helix-4b; H4b) in which R167, D171, T175 and I179 were mutated simultaneously. After four rounds of selection we isolated a pentamutant (R167N/D171S/E174S/F176Y/I179T, called H4b.4A) that bound about eightfold tighter than wild-type hGH. In a separate library (called helix-1, H1) residues F10, M14, H18 and H21, were randomly mutated. After 4 rounds of selection we isolated a tetramutant (F10A/M14W/H18D/H21N, called H1.4B) that bound threefold tighter than wild-type hGH.

Here, the phage selection studies were extended to the four contact residues in minihelix-1 (K41, Y42, L45 and Q46). These residues were randomized and representative clones were sequenced after two to seven rounds of binding selection (Table 2). Some residues were strongly selected at given positions compared to what was expected from a random distribution in the starting library. For example, about 35% of the clones contained a Q46W mutation. This was 7.6 standard deviation units ( $7.6\sigma$ ) above a random chance occurrence for Trp in the library. This way of scoring the pool of selectants establishes a consensus sequence by accounting for the expected codon bias and sampling statistics. Using this criterion there was a mild preference for K41R ( $3.7\sigma$ ), a slight preference for Y42R ( $2\sigma$ ) or Y42Q ( $2\sigma$ ), a strong preference for L45W ( $4.8\sigma$ ) or L45 ( $4.5\sigma$ ) and a stronger preference for Q46W ( $7.6\sigma$ ).

**Table 2**  
Consensus residues identified after sorting  
hGH-phagemid libraries

Residue	$P_e$	$\sigma_n$	$P_f$	$\frac{P_f - P_e}{\sigma_n}$
<b>Minihelix-1</b>				
K41 R	0.094	0.071	0.35	3.7
F	0.031	0.042	0.12	2.0
Y42 R	0.094	0.071	0.24	2.0
Q	0.031	0.042	0.18	2.0
L45 W	0.031	0.042	0.24	4.8
L	0.094	0.071	0.41	4.5
Q46 W	0.031	0.042	0.35	7.6
F	0.031	0.042	0.12	2.0
Y	0.031	0.042	0.12	2.0
<b>Loop-1</b>				
F54 P	0.062	0.047	0.73	14.1
E56 D	0.031	0.034	0.19	4.7
W	0.031	0.034	0.19	4.7
Y	0.031	0.034	0.12	2.5
I58 I	0.031	0.034	0.31	8.1
V	0.062	0.047	0.23	3.5
R64 K	0.031	0.034	0.81	22.8
<b>Combinatorial (helix-1/helix-4b)</b>				
F10 A	0.062	0.03	0.41	12.0
F	0.031	0.02	0.25	10.4
H	0.031	0.02	0.16	6.2
M14 W	0.031	0.02	0.26	11.1
S	0.094	0.04	0.26	4.8
Y	0.031	0.02	0.09	2.7
N	0.031	0.02	0.09	2.7
H	0.031	0.02	0.07	2.0
H18 D	0.031	0.02	0.43	18.8
F	0.031	0.02	0.12	4.1
N	0.031	0.02	0.10	3.4
H21 N	0.031	0.02	0.46	20.2
H	0.031	0.02	0.13	4.8
R167 N	0.031	0.02	0.63	25.6
K	0.031	0.02	0.13	4.1
D171 S	0.094	0.04	0.64	14.1
D	0.031	0.02	0.14	4.8
N	0.031	0.02	0.13	4.1
T175 T	0.062	0.03	1.0	29.1
I179 T	0.062	0.03	0.66	18.6
N	0.031	0.02	0.13	4.1

The most frequently occurring residues from phage-displayed libraries are shown, based on fractional representation ( $P_f$ ) among all clones isolated after 2 to 7 rounds of binding selection. Expected frequencies ( $P_e$ ) were calculated from the number of NNS codons for each amino acid theoretically in the starting library. Standard deviations ( $\sigma_n$ ) were calculated as  $\sigma_n = [P_e(1 - P_e)/n]^{1/2}$ . Only residues for which the fraction found exceeded the fraction expected by at least  $2\sigma_n$  are shown (i.e.  $[(P_f - P_e)/\sigma_n] > 2$ ). For the minihelix-1 library,  $n=17$  sequences; loop-1 library,  $n=26$ ; combinatorial library (helix-1/helix-4b),  $n=68$ . Twelve of the sequences from the combinatorial library contained a wild-type helix-4 region because of religation of the helix-1 vector (see Material and Methods).

A second library (called loop-1) was constructed in which F54, E56, I58 and R64 were randomly mutated. Alanine replacements at any of these residues caused a 4 to 20-fold reduction in affinity (Fig. 1(a)). Despite the fact that R64 is the only one of this group to make direct contact with the receptor (Fig. (b)), all positions showed a moderate to very strong preference for a residue that was usually different from the wild-type. R64K was the

**Table 3**  
Binding data for hGH variants isolated from the  
minihelix-1 (M1) or the loop-1 (L1) libraries

Clone	Residue position				$K_d$ (pM)	$K_d$ (hGH)
	41	42	45	46		$K_d$ (mut)
Minihelix-1						
hGH	K	Y	L	Q	340	1
M1.3A	V	S	L	W	190±26	1.8
M1.3B	L	R	L	W	190±23	1.8
M1.3C	F	R	L	Y	160±23	2.2
M1.3D	V	F	L	R	150±19	2.3
M1.3E	A	I	Q	W	ND	ND
M1.3F	L	Y	V	R	ND	ND
M1.3G	Y	W	G	Y	ND	ND
M1.3H	F	L	V	L	ND	ND
M1.5A	G	T	W	T	270±80	1.3
M1.5B	I	H	W	W	76±29	4.5
M1.5C	R	R	L	F	ND	ND
M1.5D	M	R	W	R	ND	ND†
M1.5E	R	T	A	V	ND	ND
M1.7A	R	Q	L	W	140±20	2.4
M1.7B	R	Q	L	W	140±20	2.4†
M1.7C	R	T	A	V	ND	ND
M1.7D	R	S	W	F	ND	ND
Consensus:	R	R	W	W		
		Q	L			

Clone	Residue position				$K_d$ (pM)	$K_d$ (hGH)
	54	56	58	64		$K_d$ (mut)
Loop-1						
hGH	F	E	I	R	340	1
L1.3A	P	D	T	R	210±110	1.6
L1.3B	P	Y	I	K	170±30	2.0
L1.3C	H	W	L	K	83±25	4.2
L1.3D	M	R	L	K	ND	ND†
L1.4A	G	W	V	R	660±140	0.50
L1.4B	F	W	V	R	630±120	0.53
L1.4C	S	H	L	K	620±120	0.56§
L1.4D	P	W	L	R	520±100	0.67
L1.4E	P	L	D	K	460±100	0.74
L1.4F	P	T	V	K	250±40	1.4
L1.4G	P	Y	I	K	170±30	2.0†
L1.4H	P	L	Q	K	120±30	2.8
L1.4I	P	D	T	K	61±8	5.6
L1.4J	P	T	P	K	ND	ND
L1.4K	P	A	L	K	ND	ND
L1.4L	P	C	I	K	ND	ND
L1.6A	R	D	I	R	350±250	1.0
L1.6B	P	T	V	K	250±40	1.4†
L1.6C	P	D	I	K	180±40	1.9
L1.6D	P	Y	I	K	170±30	2.0†
L1.6E	P	E	I	K	73±16	4.8
L1.6F	P	E	I	K	73±16	4.8†
L1.6G	P	D	T	K	61±8	5.6†
L1.6H	E	W	V	K	ND	ND
L1.6I	P	M	V	K	ND	ND
L1.6J	P	L	Q	K	ND	ND
Consensus:	P	D	I	K		
		W				

The nomenclature indicates the genealogy of the variant by designating its parental library (M1. or L1.) followed by the number of enrichment cycles after which it was isolated, and a letter for each sequence. Affinity constants were measured by RIA (Cunningham & Wells, 1989). The increase in affinity over hGH for binding hGHbp is shown as  $K_d(\text{hGH})/K_d(\text{mutant})$ . Some clones were not analyzed (ND). Identical affinities were assumed for equivalent variants (†). Clones with spurious mutations are indicated (E65V†; S57Y§; N47Y†; P48S||).



most preferred (found in 81% of the clones); it is known that R64K alone causes a ~threefold improvement in binding affinity (Cunningham *et al.*, 1990). After this the order of preference was F54P > I58 > E56D or W.

We analyzed the binding affinities for 27 of these mutants by expressing the free hormone in a non-suppressor host which terminates translation at the amber codon at the end of hGH and the start of the gene III domain (Lowman *et al.*, 1991a). All eight mutants tested after three to seven rounds of binding selection from the minihelix-1 library had affinities greater than wild-type hGH (Table 3). The best was K41I/Y42H/L45W/Q46W which was 4.5-fold improved in affinity over hGH. This DNA sequence is expected to occur randomly at a frequency of only one in a million clones. The fact that it was efficiently rescued demonstrates the power of the affinity selection. Of the 19 loop-1 selectants tested 13 had affinities greater than hGH. The best isolates were F54P/R64K and F54P/E56D/I58T/R64K which were both ~fivefold improved over wild-type.

#### (c) Improving affinity using additivity principles

According to additivity principles, mutations in non-interacting parts of a protein should combine to give simple additive changes in the free energy of binding (for review, see Wells, 1990). We therefore sought to improve hGH binding through site-1 by combining the substitutions isolated from phage display libraries (Fig. 2). The three tightest-binding variants of hGH from the H1 library (F10H/M14G/H18N/H21N, called H1.4A; F10A/M14W/H18D/H21N, called H1.4B; and M14S/H18F/H21L, called H1.6B) were joined to each of the three tightest binding variants found in the H4b library (R167N/D171S/E174S/F176Y/I179T, called H4b.4A; R167/D171S/E174S/F176Y, called H4b.6E; and R167N/D171N/E174S/F176Y/I179T, called H4b.6B). All variant proteins were obtained in yields approaching that of wild-type hGH except for those containing the H41.1 variant. The H1.4A variant alone or combined with mutations from the H4b library migrated as dimers ( $M_r$  ~44,000) in non-reducing SDS-PAGE and as monomers ( $M_r$  ~22,000) when reduced. Although these proteins did not contain an additional Cys residue, disulfide exchange could occur if they formed non-covalent dimers. hGH is known to form a weak dimeric complex involving residues in helices 1 and 4 (Cunningham *et al.*, 1991b). Because these proteins formed disulfide dimers we did not pursue them further. The H1.6B variant also generated a disulfide dimer, but when it was combined with the three H4b mutants no dimer was evident.

All the H1/H4b recombinants analyzed showed cumulative increases in affinity over the parental components (Table 4). The H1.4B/H4b.4A variant had the greatest affinity which was 34-fold tighter than wild-type hGH. The tightest-binding mutant from the M1 library (K41I/Y42H/L45W/Q46W,

**Table 4**  
Dissociation constants for recombinant hGH variants measured by RIA as described in Materials and Methods

Variant name	$K_d$ (pM)	$\frac{K_d(\text{hGH})}{K_d(\text{variant})}$
hGH	$340 \pm 50$	1
H1.4B†	$100 \pm 30$	3.4
H1.6B†	$680 \pm 190$	0.5
H4b.4A†	$40 \pm 20$	8.5
H4b.6E†	$40 \pm 20$	8.5
H4b.6B†	$60 \pm 30$	5.7
H1.4B/H4b.4A	$10 \pm 3$	34
H1.6B/H4b.4A	$11 \pm 3$	31
H1.6B/H4b.6E	$14 \pm 8$	24
H1.4B/H4b.6B	$16 \pm 5$	21
H1.6B/H4b.6B	$21 \pm 11$	16
M1.5B/L1.6E	$7.9 \pm 2.4$	43
H1.4B/H4b.4A/M1.5B/L1.6E	$0.9 \pm 0.3$	380

The fold improvement in binding affinity is expressed as  $K_d(\text{hGH})/K_d(\text{variant})$ . Some affinities(†) are from Lowman *et al.* (1991a). Variants are designated by a name indicating the library (e.g. H1 or H4b), the number of enrichment cycles, and a letter for each mutant. Individual variants from each library are as follows: H1.4B = (F10A/M14W/H18D/H21N), H1.6B = (M14S/H18F/H21L), H4b.4A = (R167N/D171S/E174S/F176Y/I179T), H4b.6E = (R167E/D171S/E174S/F176Y), H4b.6B = (R167N/D171N/E174S/F176Y/I179T), M1.5B = (K41I/Y42H/L45W/Q46W), and L1.6E = (F54P/R64K).

called M1.5B) and one of the tightest from the L1 library (F54P/R64K, called L1.6E) were combined to produce the hexamutant, M1.5B/L1.6E, whose affinity was 43-fold higher than wild-type hGH. This was put together with the H1.4B/H4b.4A recombinant to yield H1.4B/H4b.4A/M1.5B/L1.6E which bound 380-fold tighter than wild-type hGH. This variant retained the wild-type residue at only 5 of the 20 positions randomized (E56, I58, K172, T175, R178).

From the product of the improvements in affinity by the individual components we expected this variant to bind about 1500-fold tighter than hGH. The fact that the expected and observed values do not agree precisely suggests that on interactions between the mutated residues affect the affinity. Nonetheless, when one compares the actual *versus* the expected decrease in the free energy of binding ( $-3.6$  *versus*  $-4.3$  kcal/mol, respectively) the parity of these values is comparable to additivity experiments reported for many other proteins (Wells, 1990).

#### (d) Combinatorial libraries of hGH

In some cases we have observed very non-additive effects when mutations were introduced at neighboring residues (e.g. F176Y/E174S; Lowman *et al.*, 1991a). Therefore, because some side-chains from helix-1 can potentially contact residues in helix-4 we investigated a combinatorial approach (Huse *et al.*, 1989; Clackson *et al.* 1991) to sorting mutants derived from the H1 and H4b libraries. Independent binding selections were carried out on

**Table 5**  
hGH mutants isolated after binding selection of combinatorial libraries  
(see text for details)

Clone	P	Helix 1				Helix 4b				K <sub>d</sub> (pM)	K <sub>d</sub> (hGH) K <sub>d</sub> (mut)
		F 10	M 14	H 18	H 21	R 167	D 171	T 175	I 179		
(H1.0/H4b.0).4											
A	0.60	H	G	N	N	N	S	T	N	ND	6.8†,
B	0.40	A	N	D	A	N	N	T	N	50 ± 40	
(H1.2/H4b.2).2											
A	0.14	F	S	F	G	H	S	T	T	ND	4.6
B	0.14	H	Q	T	S	A	D	T	T	ND	
C	0.14	H	G	N	N	N	A	T	T	ND	
D	0.14	F	S	F	L	S	D	T	T	ND	
E	0.14	A	S	T	N					ND	
F	0.14	Q	Y	N	N	H	S	T	T	74 ± 30	
G	0.14	W	G	S	S	---				ND	
(H1.2/H4b.2).2 (repeat)											
H	0.13	F	L	S	S	K	N	T	V	ND	2.1
I	0.13	W	N	N	S	H	S	T	T	160 ± 70	
J	0.13	A	N	A	S	N	S	T	T	ND	
K	0.13	P	S	D	N		---	---	---	ND	
L	0.13	H	G	N	N	N	N	T	T	ND	
M	0.13	F	S	T	G				---	ND	
N	0.13	M	T	S	N	Q	S	T	T	ND	
O	0.13	F	S	F	L	T	S	T	T	ND	
(H1.2/H4b.2).4											
A	0.17	A	W	D	N	---	---	---		100 ± 30	3.3†
B	0.17	A	W	D	N	H	S	T	N	ND	
C	0.17	M	Q	M	N	N	S	T	T	NE§	0.4†
D	0.17	H	Y	D	H	R	D	T	T	ND	
E	0.17	L	N	S	H					820 ± 200	
F	0.17	L	N	S	H	T	S	T	T	34 ± 19	
(H1.2/H4b.2).6											
A	0.38	A	W	D	N			---	---	100 ± 30	3.3†
B	0.13	A	W	D	N	N	S	T	S	ND	
C	0.13	A	W	D	N	K	D	T	T	ND	5.0†,‡
D	0.13	A	T	S	N	N	S	T	T	ND	
E	0.13	M	A	D	N	N	S	T	T	68 ± 46	
F	0.13	H	Y	D	H	N	S	T	T	ND	
(H1.2/H4b.2).6 (repeat)											
G	0.38	A	H	A	S	N	S	T	T	ND	28
H	0.25	F	S	L	A	N	S	T	I	ND	
I	0.13	H	Y	D	H	Y	S	T	S	ND	
J	0.13	V	L	D	H	N	S	T	T	ND	
K	0.13	A	W	D	N	N	N	T	I	ND <sup>c</sup>	
(H1.2/H4b.2).7											
A	0.33	A	W	D	N	N	A	T	T	12 ± 6	3.3†
B	0.17	A	W	D	N	---	---	---	---	100 ± 30	
C	0.08	A	W	D	N	N	S	T	N	ND	21
D	0.08	A	W	D	N	R	N	T	T	ND	
E	0.08	A	W	D	N	K	S	T	S	ND	
F	0.08	F	S	T	G					ND	
G	0.08	I	Q	E	H	N	S	T	T	16 ± 10	
H	0.08	H	Y	D	H	N	S	T	T	ND	
(H1.2/H4b.2).7 (repeat)											
I	0.50	F	S	L	A	N	S	T	V	32 ± 5	11
J	0.25	A	H	A	S	N	S	T	T	ND	
K	0.13	A	W	D	N	A	N	T	T	ND	
L	0.13	H	Y	D	H	Y	S	T	S	ND	
(H1.4/H4b.4).4											
A	0.67	F	S	F	L	K	D	T	T	150 ± 70	2.3†
B	0.17	F	S	F	L	N	S	T	T	11 ± 3	31.0†
C	0.17	M	A	D	N	N	S	T	T	68 ± 46	5.0†,‡

All variants contain (E1745S/F176Y), except for those with the wild-type helix 4 sequence (—), which were non-recombinants. Libraries are designated by the original helix library followed by the number of independent enrichment cycles. For example, (H1.2/H4b.2).4 indicates that helix-1 and helix-4b libraries were sorted independently for 2 rounds (H1.2/H4b.2). DNA pools were recombined, and the combinatorial library was sorted 4 more times (.4). "repeat" indicates that the library was reconstructed and sorted independently. The clone letter identifies the specific mutant. P, indicates the fractional occurrence among sequenced clones. Some affinities are from Lowman *et al.* (1991a); equivalent variants are assumed to have identical affinities(†). Several variants appeared as >10% disulfide dimers(‡). One clone contained an amber (TAG = Gln in *SupE* strains) codon(§), one contained a spurious mutation, E174N(¶), and one(||) contained 2 mutations (L15R/K168R). Some variants were not expressed (NE) or not analyzed (ND).

the H1 and H4b libraries for 0, 2 or 4 cycles. DNA from the H1 pool was ligated together with DNA from the H4b pool that was sorted for binding to the hGHbp for the same number of rounds. The three combinatorial libraries (H1.0/H4b.0, H1.2/H4b.2 and H1.4/H4b.4) were then sorted an additional two to seven cycles and 68 clones were sequenced (Table 5).

Overall, the sequences of the highest affinity variants isolated from any of the combinatorial libraries resembled those assembled by independent sorting of the H1 and H4b libraries (Lowman *et al.*, 1991a) and recombination described above. For example, the highest affinity mutants isolated previously from the H1 library were F10A/M14W/H18D/H21N (H1.4B) and F10H/M14G/H18N/H21N (H1.4A) which bound about 3.3-fold and 2.4-fold tighter than wild-type hGH, respectively. The H1.4A sequence was recovered in 60% of the clones from combinatorial library (H1.0/H4b.0).4, and in 13% of the clones isolated in early rounds of sorting from combinatorial library H1.2/H4b.2. The H1.4B sequence predominated in later rounds of sorting the H1.2/H4b.2 library. Most of these were independent clones (not siblings or contaminants) because they had different DNA sequences and usually differed in the mutations selected in helix 4.

Similar results were obtained with selectants in helix 4. When the H4b library was independently sorted, many sequences containing R167N, D171S or N, T175 and I179T were selected (Lowman *et al.*, 1991a). These were the same residues that tended to be selected in any of the combinatorial libraries. In fact, one of the best mutants previously isolated (R167N/D171S/T175/I179T) was commonly isolated by combinatorial sorting and predominated in the later rounds.

Some sequences sorted by combinatorial means were very different from ones selected from the two independent libraries, perhaps for statistical reasons. For example, the H1 and H4b libraries contained about  $10^6$  different DNA sequences, and if combined (without pre-selection) would contain  $10^{12}$  possible combinations. Transformation efficiencies limit the sampling size to  $\sim 10^7$  independent clones. Thus, the selection of the same sequences is remarkable given the high diversity of sequences possible in these libraries and the mild increases in affinity being selected for.

We measured the affinities for a number of the isolates from the combinatorial libraries (Table 5). All had improved binding affinity (2 to 31-fold) compared to wild-type hGH. Most were improved over E174S/F176Y which was present in all the starting clones, and independently caused a 5.6-fold increase in affinity over wild-type hGH (Lowman *et al.*, 1991a). The highest affinity variants were generally isolated from later rounds of sorting and were highly abundant in those pools. For example, one of the highest affinity mutants we tested was clone (H1.2/H4b.2).7A which was isolated after seven rounds of sorting of H1.2/H4b.2 library. This isolate represented a third of the clones in that pool.

Remarkably, this clone is identical to the H1.4B/H4b.4A variant (Table 4) except that instead of D171S it contained the conservative substitution, D171A. Not surprisingly, these variants bound with virtually the same affinities (12 pM and 10 pM, respectively. In this case, the combinatorial and additive strategies yield comparable solutions for successful optimization of affinity.

#### (e) Testing the importance of individual side-chains in affinity maturation

To evaluate the contribution of mutated residues to the binding affinity we introduced them into wild-type hGH, or converted them back to the wild-type residue in some of the affinity improved variants (Table 6). The K41I/Y42H/L45W/Q46W tetramutant bound 4.5-fold tighter than wild-type hGH. Yet, each of the single mutants in hGH caused 1.7 to 2.5-fold reductions in affinity. This indicates that the combination of mutations at this site is critical for the affinity improvements. These residues lie on adjacent positions on one face of the minihelix-1. Non-additive effects are well documented for neighboring residues in other proteins (Wells, 1990) as well for hGH (Lowman *et al.*, 1991a).

Affinity improvements caused by substitutions in the H1.4B/H4b.4A variant were tested by mutating them back to the wild-type residue either individually or in adjacent pairs (Table 6). This showed that seven of the nine substitutions contribute only small or virtually undetectable improvements in binding (1.1 to 1.7-fold). Even the most dominant effect, F176Y, imparts only a 4.6-fold improvement in binding affinity. Nonetheless, the product of

**Table 6**  
Importance of individual side-chains in various hGH mutants (described in Tables 3 and 4) for binding to the hGHbp

	$K_d$ (pM)	$\frac{K_d(\text{mut})}{K_d(\text{hGH})}$	$\frac{K_d(\text{mut})}{K_d(\text{H1.4B/H4b.4A})}$
Point mutants in wild-type background			
hGH (wild-type)	340 ± 50	1.0	—
K41I†	580 ± 140	1.7	—
Y42H†	860 ± 50	2.5	—
L45W†	722 ± 60	2.1	—
Q46W†	780 ± 100	2.3	—
Revertants in H1.4B/H4b.4A background			
(H1.4B/H4b.4A)	10 ± 3	—	1.0
D18H/N21H	12 ± 9	—	1.1
A10F/W14M	13 ± 5	—	1.2
A10F†	13 ± 4	—	1.3
N167R/S171D	17 ± 8	—	1.6
T179I	18 ± 9	—	1.7
Y176F	49 ± 21	—	4.6

Binding affinities were measured by BIAcore™(†) or by RIA and normalized to the RIA value for hGH determined previously (Cunningham & Wells, 1989). The fold decrease in affinity is expressed as  $K_d(\text{revertant})/K_d(\text{parent})$ , where the parent is the background used for mutagenesis.



these effects in the octamutant, F10A/M14W/H18D/H21N/R167N/D171S/F176Y/I179T (not tested here), predicts a 16-fold improvement in affinity *versus* wild-type hGH. This compares to the 34-fold enhancement measured for the H1.4B/H4b.4A variant. This simple additivity calculation is in reasonable agreement with experiment considering that S174E, which was not tested in the H1.4B/H4b.4A context, was expected to enhance affinity an additional two- to threefold. Thus, the affinity improvements come by a series of small effects that combine to yield large affinity enhancements.

(f) *Effects of affinity maturation on the kinetics of binding*

To better understand the molecular basis for affinity improvements selected here, we used the BIAcore™ to measure the kinetics of binding to the hGHp (Table 7). In general, as we increased the affinity from wild-type hGH, the off-rate decreased with little change in on-rate. In fact, in going from wild-type to the highest affinity mutant, H1.4B/H4b.4A/M1.5B/L1.6E, there was more than a 60-fold decrease in the off-rate and only a fourfold increase in the on-rate. (For this mutant the off-rate was too slow to measure accurately, but if we calculate it from the on-rate and the  $K_d$  measured by RIA, the off-rate would be 100-fold slower than wild-type hGH.) The fact that off-rate is most affected among the phage selectants suggests that the sorting was performed under conditions approaching equilibrium. We had previously recruited the hGH binding site into hPL which differs in sequence by ~15% from hGH and binds ~2000-fold weaker (Lowman *et al.*, 1991b). The recruited hPL variant has kinetic parameters for binding that are similar to hGH (Table 7). Relative to wild-type hPL, this mutant shows much larger improvements in off-rate (~100-fold) compared to on-rate (~10-fold).

**Table 7**  
*Binding kinetics of hGH variants*

Mutant	$k_{on}/10^4$ $M^{-1} s^{-1}$	$k_{off}/10^{-5}$ $s^{-1}$	$K_d(nM)$	$K_d(hGH)$ $K_d(mut)$
hPL	3.2	6000†	1800	0.0006
hPL variant	43	49	1.1	0.79
hGH	24	34	1.4	1
M1.5B	13	6.9	0.52	2.7
L1.6E	21	6.6	0.31	4.5
M1.5B/L1.6E	36	5.1	0.14	10
H1.4B/H4b.4A	20	3.0	0.15	9.3
H1.4B/H4b.4A/ M1.5B/L1.6E	98	≤0.6	≤0.006	≥230

Kinetic constants were determined using a BIAcore™ in which the (S201C)hGHbp was immobilized on the biosensor chip in PBS buffer ±0.05% Tween-20. The  $K_d$  is calculated from  $k_{off}/k_{on}$ , except for hPL, for which  $k_{on}$  and  $K_d$  were measured and  $k_{off}$  calculated(†). The ratio of  $K_d$  values indicates the fold increase in binding affinity *versus* wt-hGH. hGH mutants are described in Table 4 and the hPL variant contains V41/D56E/M64R/E174A/M179I.

#### 4. Discussion

(a) *Monovalent phage display is a powerful selection procedure*

We randomly mutated regions of hGH which were thought to be important either because they were in contact with the receptor or because when converted to alanine they affected binding affinity. Thus, an average random mutant from these libraries should be dramatically reduced in binding affinity from wild-type hGH. Yet clones isolated after only a few rounds of selection bound with similar and often higher affinity than wild-type hGH. These usually exhibited consensus sequences that were different from the wild-type residue (Table 2).

Small improvements in affinity led to rapid and almost exclusive convergence in these libraries. For example, the R64K mutant binds only ~three times tighter than wild-type hGH (Cunningham *et al.*, 1990). After just three cycles of binding selection, R64K dominated the library (Table 3). Similarly, I179T contributed only a 1.7-fold improvement in affinity (Table 6). However, when sorted separately in the H4b library (Lowman *et al.*, 1991a) or combinatorially with mutants in H1 (Tables 2, 5), I179T was selected almost exclusively. Strong selection for these slight improvements in affinity emphasize the power of this technique for rescuing the highest affinity variants in the pool.

The phage display method has important limitations also (for reviews, see Smith, 1991; Wells & Lowman, 1992). For example, not all variants will be displayed on the phage because they may be either digested by proteases, aggregated, misfolded or blocked in secretion or assembly on phage. Although a strong bias against particular DNA sequences is unlikely, there is a clear selection against Cys-containing mutants. This has been previously noted for mutants of hGH (Lowman *et al.*, 1991a) as well as other proteins (Matthews & Wells, 1993) and may be caused by an odd thiol fouling the disulfide bonding scheme in the selected protein binding domain, or in the gene III protein itself which also contains disulfides (van Wezenbeck *et al.*, 1980). Limits in transfection efficiency restrict the number of possible sequences one can screen so one should be selective about the codons randomly mutated.

(b) *Simple additivity dominates affinity maturation*

Additivity principles (for a review, see Wells, 1990) suggest that when residues are in close contact, the sum of the free energy effects for the components may not equal the combined mutant because the mutant side-chains can interact intramolecularly, according to the following equation:

$$\Delta\Delta G_{AB} = \Delta\Delta G_A + \Delta\Delta G_B + \Delta\Delta G_I \quad (1)$$

where  $\Delta\Delta G_{AB}$  is the change in free energy for the double mutant,  $\Delta\Delta G_A$  and  $\Delta\Delta G_B$  are the changes in free energy for the two single mutants, respectively, and  $\Delta\Delta G_I$  is the change in the interaction free



energy between the two residues. When the residues do not contact, they act independently, and thus the component residues should contribute simple additive effects on the binding free energy (i.e.  $\Delta\Delta G_i = 0$ ). Studies of multiple mutants of hGH, hPRL and hPL have shown that this relationship holds for many substitutions throughout the hGH molecule (Cunningham *et al.*, 1990; Cunningham & Wells, 1991; Lowman *et al.*, 1991b).

Residues on the same face of an  $\alpha$ -helix were mutated simultaneously because some side-chains displayed from them could contact one or more of the others. In fact, several examples of non-additivity (i.e.  $\Delta\Delta G_i$  not zero) were seen within a given library. The K41I/Y42H/L45W/Q46W mutant isolated from the M1 library bound 4.5-fold tighter than wild-type hGH (Table 3). However, when each of these residues were installed into wild-type hGH they caused 1.7 to 2.5-fold reductions in affinity (Table 6). A similar finding was reported for the F176Y mutant in the context of E174S/F176Y (Lowman *et al.*, 1991a). Despite these examples, simple additivity within each library was usually found. For example, F54P/E56D/I58T/R64K bound 3.5-fold tighter than the triple mutant, F54P/E56D/I58T (Table 3). The R64K mutant binds ~threefold tighter than wild-type hGH (Cunningham *et al.*, 1990) which can account for the difference between the tetra-mutant and the triple mutant. As another example, the F54P/R64K double mutant bound 2.5-fold tighter than the triple mutant F54P/E56D/R64K (Table 3). The D56E mutant in hPL caused a 2.4-fold improvement in binding to the hGHbp (Lowman *et al.*, 1991b). Further evidence for simple additivity is that residues at many positions of each library showed a consensus that was independent of variation at the neighboring residues. This argues that their impact on affinity was generally independent of the other residues that were mutated. Lastly, the highest affinity variants selected from the H1 and H4a libraries were virtually the same whether sorted independently or sorted together (Table 5). This occurred even though some side-chains from the helix-1 library interact with those in helix-4b. For example, His21 makes a hydrogen bond with Asp171.

Combining the affinity enhanced variants from separate libraries generally produced simple additive improvements in binding free energy (Table 4). Thus, the modest gains in affinity (3 to 6-fold) from each of the mutant libraries accumulated to produce a combined variant with a 380-fold improved affinity. Each of the 15 mutant side-chains imparts only a small gain in affinity as evidenced by the small or sometimes insignificant reductions in affinity when various side-chains in the H1.4B/H4b.4A mutant are converted back to the wild-type (Table 6).

The small gains in affinity produced by each mutant side-chain can sometimes be rationalized from the structure of the bound complex and modeled side-chain replacements. For example,

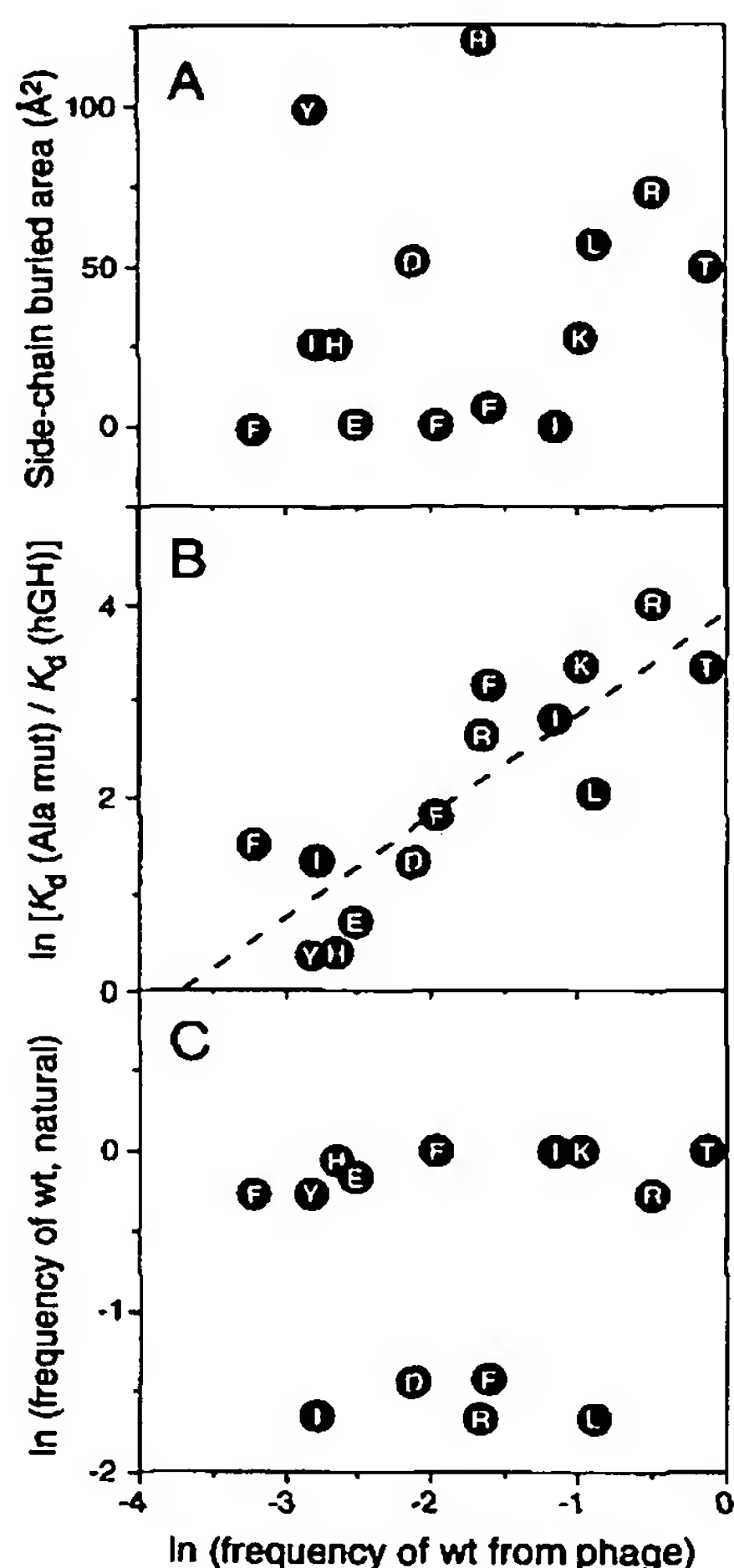
I179T may form a hydrogen bond to the main-chain amide of Gly168 in the receptor. The salt-bridge between Arg64 of hGH and Asp164 of the receptor may be improved by substituting the shorter Lys64 side-chain. This could reduce a repulsive interaction between Arg64 ( $N_{H1}$ ) and receptor Lys167 ( $N_2$ ), less than 4.5 Å away. Of course without the structure of the bound and free components we cannot attribute the enhanced affinities to improvements in the packing of the bound complex or destabilization of the free components. Recently, the structure of the free affinity optimized mutant containing 15 substitutions has been solved (Ultsch *et al.*, 1993). The overall structure is very close to that of wild-type hGH in the bound hGH(hGHbp)<sub>2</sub> complex.

The affinity enhancements that occurred for the hGH variants compared to wild-type hGH, or for the recruited hPL variant compared to wild-type hPL, were caused largely by decreasing the off-rate, not increasing the on-rate (Table 7). This is not surprising given the previous finding that the largest reduction caused by an alanine substitution in on-rate is only threefold, whereas the largest reduction in off-rate is about 30-fold (Cunningham & Wells, 1993). We have suggested this is because there are many pathways leading to association, but there is likely to be only one bound complex. Analysis of a mutant antibody that is improved in affinity similarly shows the off-rate is decreased with little effect upon on-rate (Marks *et al.*, 1992).

Northrup & Erickson (1992) have suggested that association rate is controlled by diffusion and long-lived collisions that result from Brownian dynamic effects. While these factors are largely independent of side-chain substitutions, we estimate that positively charged side-chains at the site 1 epitope can contribute up to a factor of 20 to the association rate (Cunningham & Wells, 1993). Consistent with these electrostatic effects, the recruited hPL variant (V41I/D56E/M64R/E174A/M179I), which associates tenfold faster than wild-type hPL, contains M64R and E174A (Table 7). These make the hormone more electropositive at positions where alanine substitutions in hGH are known to affect on-rate by about twofold each (Cunningham & Wells, 1993). Furthermore, the hGH mutant, K41I/Y42H/L45W/Q46W, is reduced almost twofold in association rate which may be due to the K41I substitution. The on-rate effects are more subtle and accumulate in a somewhat more complex fashion than the off-rate effects. For example, most of the component mutants used to derive the H1.4B/H4b.4A/M1.5B/L1.6E variant (whose on-rate is about 4-fold higher than wild-type hGH) had on-rates that were less than or equal to wild-type hGH.

#### (c) Relationship of phage sorting to structural and functional epitopes

Less than half of the side-chains that become buried at site 1 by the first receptor significantly affect binding affinity when converted to alanine



**Figure 3.** Relationship between the frequency that a wild-type residue is recovered after randomization and phage selection *versus* the extent to which the side-chain is buried upon binding the receptor (A), or is reduced in binding affinity when the residue is converted to alanine (B), or the frequency that the residue is conserved in the natural family of growth hormones (C). The natural logarithm of the frequency with which wild-type hGH residues were selected was calculated from the phage sorting data shown in Table 2 and Lowman *et al.* (1991a). The log scale was chosen to scale the data and for comparison with buried surface area which has been related to binding free energy. Data for residues M14, H18, K41, Q43, R167 and E174 could not be used in the graph because the wild-type residue was not recovered at any of these positions. Wild-type residues from left to right across the X axis are F, Phe54; I, Ile179; Y, Tyr42; H, His21; E, Glu56; D, Asp171; F, Phe10; R, Arg64; F, Phe176; I, Ile58; K, Lys172; L, Leu45; R, Arg178; and T, Thr175. In A the change in solvent-accessible surface area of each side-chain upon binding was calculated as the accessible areas of the side-chain in free hGH minus that in the hGH(hGHbp)<sub>2</sub> complex (see Cunningham & Wells, 1993). The correlation coefficient ( $R^2$ ) for this plot was 0.022. In B the relative change in binding affinity from wild-type hGH for alanine substitutions is plotted as  $\ln [K_d(\text{Ala mutant})/K_d(\text{hGH})]$  to scale to free energy. These parameters are mildly correlated ( $R^2=0.71$ ), with a slope near unity ( $y=1.0 \times +3.9$ ). In C the conservation of residues among 17 natural growth hormone variants was determined from the amino acid sequences (Genbank, v.

Cunningham & Wells, 1993; Fig. 1(a)). The mini-helix-1 contact residues provide a good example of this (Table 1). The side-chains of Tyr42 and Gln46 each make more van der Waals contacts and undergo more burial than Leu45. Yet, Y42A and Q46A have much smaller effects upon binding compared to L45A. These studies suggest that functionality is not easily assessed by the extent to which a side-chain makes contact with the receptor.

Another way to evaluate this is to correlate the conservation of wild-type residues after binding selection with the extent to which they are buried by the receptor. As shown in Figure 3A, there is no correlation between these parameters ( $R^2<0.05$ ). This is also evident by comparing the structural and phage conservation epitopes shown in Figure 1(b) and (c), respectively. Binding affinity is determined by the difference in energy between the bound and free states. Clearly the X-ray structure of the complex (de Vos *et al.*, 1992) only reveals the molecular details of the bound state. Solvation and intramolecular interaction energies in the free state, which we have little current information on, can dramatically affect the energetics of binding.

(d) *Selected residues at functionally important and buried sites tend to be identical or isosteric with the wild-type*

There is a moderate correlation ( $R^2=0.71$ ) between the reduction in affinity as assessed by alanine-scanning and side-chain conservation following phage sorting (Fig. 3B; Lowman *et al.*, 1991a). This indicates that functionality determined by alanine-scanning is similar to that determined by sequence conservation after binding selection (also compare Fig. 1(a) and (c)). There is not a good correlation between the frequency of conservation of given residues among natural variant growth hormones and conservation following binding selection (Fig. 3C). We have argued that natural variation occurs in the corresponding hormone receptors to complement mutations in the hormone (Lowman *et al.*, 1991a). Also, hGH binds to the prolactin receptor which is coded for separately from the hGH receptor. Thus, in nature the functional constraints on growth hormone are not fixed by a single and non-adaptable target receptor as they are by the *in vitro* binding selection.

We find that selected residues at the most functionally important and highly buried sites, either at the interface or in the hormone itself, tend to be retained as the wild-type residue or a close homolog. For example, all five of the residues that are most conserved as the wild-type after extensive phage

75, Feb., 1993) from monkey, pig, elephant, hamster, whale, alpaca, fox, horse, sheep, rat, turtle, chicken, mink, cow, salmon, frog and trout, as well as human placental lactogen, hGH(20K) and hGH-V. The correlation coefficient for this plot was  $R^2=0.005$ .



sorting (E56, I58, K172, T175 and R178) are completely buried in the complex; converting them to alanine caused 4 to 60-fold reductions in affinity (Cunningham & Wells, 1989). When substitutions were tolerated at these positions they were typically similar to the wild-type residue. For example, the highest affinity selectants contained either Asp or Glu at position 56,  $\beta$ -branched residues at position 58, Lys or Arg at 172, Thr or Ser at 175, and Lys or Arg at 178 (Table 3; Lowman *et al.*, 1991a).

There is another group of functional residues that become highly buried upon receptor binding (K41, L45, R64, D171 and I179). When these were randomized we found improved substitutes that tended to be similar in character to the wild-type residue. For example, Lys41 was often replaced with Arg; Leu45 was substituted with large hydrophobic side-chains; Arg64 was most frequently substituted by Lys; Asp171 was optimally replaced by Asn and sometimes Ser; Ile179 was usually substituted by  $\beta$ -branched residues (Tables 2, 3, 5; Lowman *et al.*, 1991a). Thus, improvements can be made at functionally important residues buried at the interface. These substitutions tend to be toward an isosteric side-chain or one of similar chemical character.

Two of the residues that were randomized (His18 and Glu174) had enhanced binding affinity when converted to alanine and were completely buried in the complex. These almost always sorted to something smaller than the wild-type residue. For example, the preferred substitution for His18 was Asp or Asn, and for Glu174 was Ala, Ser or Thr (Lowman *et al.*, 1991a). This suggests that packing at these positions is energetically unfavorable.

Another class of residues (H21, Y42, Q46 and R167) are highly buried at the interface, but have little or no effect on binding affinity when converted to alanine. These residues rarely sort back to the wild-type residue. For example, His21 tended to sort to Asn; Tyr42 often came back as Arg or Gln; Gln46 converted to Trp and Arg167 often sorted to Asn (Tables 2, 3, 5; Lowman *et al.*, 1991a). Despite the consensus found at these buried residues, the affinity enhancements made from them were very small (Table 6). Thus, it appeared more difficult to obtain improvements in affinity from contact residues that were functionally inert to start with.

The last group of residues (F10, M14 and F54) are virtually buried in the folded hormone and affected binding affinity by two- to fourfold when converted to alanine, presumably by indirect structural effects. Surprisingly radical substitutions were tolerated here that showed consensus sorting (Tables 2, 3, 5; Lowman *et al.*, 1991a). For example, F54P was almost the sole solution in the L1 library. Phe54 is 84% buried in the hormone and 10 Å away from making contact with the receptor. We estimate the F54P mutant enhances affinity by a factor of  $\sim 1.6$ -fold based on the fact that the double mutant (F54P/R64K) is improved in binding by 4.8-fold (Table 3) and the R64K mutant alone enhances binding by a factor of  $\sim 3$  (Cunningham *et al.*, 1990).

Although it is possible to rationalize the general features of these mutants by combining the functional and structural data, there were always unusual mutants that came through the sorting. For example, Ile179 was almost always conservatively replaced by a  $\beta$ -branched side-chain (especially Ile or Thr), but I179S has appeared (Table 5). Similarly, Leu45 was almost always replaced by a large side-chain (Leu or Trp), but L45A was also found (Table 3). Some of these may appear because of a background of weaker affinity variants in the selected phage pools.

#### (e) Phage sorting strategies

These studies suggest guidelines for affinity maturation of binding interfaces using monovalent phage display. One cannot easily search more than five or six codons exhaustively (Lowman & Wells, 1991), therefore the library needs to be focused on residues where one can hope to improve affinity. A good starting point for targeting residues for optimization of affinity is a complete alanine-scan of the relevant interface. It is also possible to limit the codon choices (see e.g. Arkin & Youvan, 1992), and this is especially reasonable if one has detailed knowledge of the structural interface to begin with.

The residues where the most obvious improvements in affinity occurred were those that were shown by alanine-scanning to significantly affect binding. For example, we estimate the largest improvements in affinity came from R64K, E174S and F176Y. E174A was known to enhance affinity, but R64A and F176A caused large reductions in affinity. Thus, despite the fact that the most highly conserved residues in the phage sorting were those that were most important by alanine-scanning, we still found improved variants here. It is also noteworthy that all three of these residues are functionally important for binding to the prolactin receptor (Cunningham & Wells, 1991). Thus, nature may have selected these residues for their dual role, not simply for optimal binding to the hGH receptor.

Ideally, one should randomize residues that can contact each other in the same mutagenesis step so that they can covary. We saw covariance in the M1 and H4a libraries when residues were close enough to interact. Sorting libraries by combinatorial means is especially useful in situations where mutations may lead to non-additive effects. For example, if side-chain replacements cause large conformational changes, as they may in flexible loops of antibodies, combinatorial sorting would allow for improvements by searching randomly for the best combinations of mutant heavy and light chains (Huse *et al.*, 1989; Clackson *et al.*, 1991; Collet *et al.*, 1992).

In general the improvements in hGH tended to occur by simple additive effects both between libraries and within libraries. Practically, this means that one can randomize many residues independently and combine them in the end to obtain high affinity

variants. Fundamentally, it suggests that the interactions between side-chains on hGH often have little overall effect on the free energy of binding receptor.

H.L. was supported in part by NIH grant GM13560-03. We thank Parkash Jhurani, Peter Ng and Mark Vasser for oligonucleotide synthesis, Abraham de Vos for hGH(hGHbp)<sub>2</sub> crystal co-ordinates and helpful discussions, Brian Cunningham for the (S201C)hGHbp, several hGH variants and useful discussion, Lei Jin for (K41Q)hGH, Nancy McFarland for providing *E. coli* strains, Mike Covarrubias and Brad Snedecor for fermentations, Marjorie Winkler for purifying hGHbp, Mark Ultsch for providing several purified hormone variants, John Stultz for mass spectrometry, Kerrie Andow and Tom Hynes for graphics, and Tom Camerato for help with DNA sequencing.

### References

- Arzin, A. P. & Youvan, D. C. (1992). Optimizing nucleotide mixtures to encode specific subsets of amino acids for semi-random mutagenesis. *Bio/Technology*, **10**, 297-300.
- Barnard, R., Bundesen, P. G., Rylatt, D. B. & Waters, M. J. (1984). Monoclonal antibodies to the rabbit liver growth hormone receptor: production and characterization. *Endocrinology*, **115**, 1805-1813.
- Bass, S., Green, R. & Wells, J. A. (1990). Hormone phage: an enrichment method for variant proteins with altered binding properties. *Proteins*, **8**, 309-314.
- Bevington, P. R. (1969). *Data Reduction and Error Analysis for the Physical Sciences*, pp. 56-65. McGraw-Hill, New York.
- Chang, C. N., Rey, M., Bochner, B., Heynecker, H. & Gray, G. (1987). High-level secretion of human growth hormone by *Escherichia coli*. *Gene*, **55**, 189-196.
- Clackson, T. M., Hoogenboom, H. R., Griffiths, A. D. & Winter, G. (1991). Making antibody fragments using phage display libraries. *Nature (London)*, **352**, 624-628.
- Collet, T. A., Roben, P., O'Kennedy, R., Barbas, C. F., III, Burton, D. R. & Lerner, R. A. (1992). A binary plasmid system for shuffling combinatorial antibody libraries. *Proc. Nat. Acad. Sci., U.S.A.* **89**, 10026-10030.
- Cunningham, B. C. & Wells, J. A. (1989). High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science*, **244**, 1081-1085.
- Cunningham, B. C. & Wells, J. A. (1991). Rational design of receptor-specific variants of human growth hormone. *Proc. Nat. Acad. Sci., U.S.A.* **88**, 3407-3411.
- Cunningham, B. C. & Wells, J. A. (1993). Comparison of a structural and functional epitope. *J. Mol. Biol.* **234**, 554-563.
- Cunningham, B. C., Henner, D. J. & Wells, J. A. (1990). Engineering human prolactin to bind to the human growth hormone receptor. *Science*, **247**, 1461-1465.
- Cunningham, B. C., Ultsch, M., de Vos, A. M., Mulkerrin, M. G., Klausner, K. R. & Wells, J. A. (1991a). Dimerization of the extracellular domain of the human growth by a single hormone molecule. *Science*, **254**, 821-825.
- Cunningham, B. C., Mulkerrin, M. G. & Wells, J. A. (1991b). Dimerization of human growth hormone by zinc. *Science*, **253**, 545-548.
- de Vos, A. M., Ultsch, M. & Kossiakoff, A. A. (1992). Human growth hormone and the extracellular domain of its receptor: crystal structure of the complex. *Science*, **255**, 306-312.
- Fuh, G., Cunningham, B. C., Fukunaga, R., Nagata, S., Goeddel, D. V. & Wells, J. A. (1992). Rational design of potent antagonists to the human growth hormone receptor. *Science*, **256**, 1677-1680.
- Huse, W. D., Sastry, L., Iverson, S. A., Kang, A. S., Alting-Mees, M., Burton, D. R., Benkovic, S. J. & Lerner, R. A. (1989). Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. *Science*, **246**, 1275-1281.
- Jin, L., Fendly, B. M. & Wells, J. A. (1992). High resolution functional analysis of antibody-antigen interactions. *J. Mol. Biol.* **226**, 851-865.
- Karlsson, R., Michaelsson, A. & Mattson, A. (1991). Kinetic analysis of monoclonal antibody-antigen interactions with a new biosensor based analytical system. *J. Immunol. Methods* **145**, 229-240.
- Kunkel, T. A., Bebenek, K. & McClary, J. (1991). Efficient site-directed mutagenesis using uracil-containing DNA. *Methods Enzymol.* **204**, 125-139.
- Löfås, S. & Johnsson, B. (1990). A novel hydrogel matrix on gold surfaces in surface plasmon resonance sensors for fast and efficient covalent immobilization of ligands. *J. Chem. Soc. Chem. Commun.* **21**, 1526-1528.
- Lowman, H. B. & Wells, J. A. (1991). Monovalent phage display: a method for selecting variant proteins from random libraries. *Methods: Companion Methods Enzymol.* **3**, 205-216.
- Lowman, H. B., Bass, S. H., Simpson, N. & Wells, J. A. (1991a). Selecting high-affinity binding proteins by monovalent phage display. *Biochemistry*, **30**, 10832-10838.
- Lowman, H. B., Cunningham, B. C. & Wells, J. A. (1991b). Mutational analysis and protein engineering of receptor-binding determinants in human placental lactogen. *J. Biol. Chem.* **266**, 10982-10988.
- Marks, J. D., Griffiths, A. D., Malmqvist, M., Clackson, T. P., Bye, J. M. & Winter, G. (1992). By-passing immunization: building high affinity human antibodies by chain shuffling. *Bio/Technology*, **10**, 779-783.
- Matthews, D. J. & Wells, J. A. (1993). Substrate phage: selection of protease substrates by monovalent phage display. *Science*, **260**, 1113-1117.
- Northrup, S. H. & Erickson, H. P. (1992). Kinetics of protein-protein association explained by Brownian dynamics computer simulation. *Proc. Nat. Acad. Sci., U.S.A.* **89**, 3338-3342.
- Olson, K. C., Fenno, J., Lin, N., Harkins, R. N., Snider, C., Kohr, W. H., Ross, M. J., Fodge, D., Prender, G. & Stebbing, N. (1981). Purified human growth hormone from *E. coli* is biologically active. *Nature (London)*, **293**, 408-411.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5463-5467.
- Smith, G. P. (1991). Surface presentation of protein epitopes using bacteriophage expression systems. *Curr. Opin. Biotechnol.* **2**, 668-673.
- Ultsch, M., Somers, W., Kossiakoff, A. A. & de Vos, A. M. (1993). The crystal structure of affinity-matured



- growth hormone at 2 Å resolution. *J. Mol. Biol.* in the press.
- van Wezenbeck, P. M. G. F., Hulsebos, T. J. M. & Shoenmakers, J. G. G. (1980). Nucleotide sequence of the filamentous bacteriophage M13 genome: comparison with phage fd. *Gene*, **11**, 129-148.
- Vieira, J. & Messing, J. (1987). Production of single-stranded plasmid DNA. *Methods Enzymol.* **153**, 3-11.
- Wells, J. A. (1990). Additivity of mutational effects in proteins. *Biochemistry*, **29**, 8509-8517.
- Wells, J. A., Vasser, M. & Powers, D. B. (1985). Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites. *Gene*, **34**, 315-323.
- Wells, J. A. & Lowman, H. B. (1992). Rapid evolution of peptide and protein binding properties *in vitro*. *Curr. Opin. Struct. Biol.* **2**, 597-604.
- Wells, J. A. & de Vos, A. M. (1993). Structure and function of human growth hormone: implications for the hematopoietins. *Annu. Rev. Biophys. Biomol. Struct.* **22**, 329-351.

Edited by P. E. Wright